Structural and biochemical investigations into c-di-GMP signaling in *E.coli*

Zinc dependent regulation of the diguanylate cyclase YdeH and characterization of PgaA and PgaB, involved in c-di-GMP controlled exopolysaccharide synthesis

INAUGURALDISSERTATION

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Franziska Zähringer
aus Deutschland

Basel, 2013

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von

Prof. Dr. Tilman Schirmer
Prof. Dr. Urs Jenal

Basel, den 21.02.2012

Prof. Dr. Martin Spiess

Contents

Abstract

In response to adverse conditions, many bacterial species can switch from a planktonic growth to a surface associated growth mode and form biofilm communities. A key factor triggering the formation of biofilms in a multitude of bacterial species is the second messenger bis-(3'-5')cyclic dimeric guanosine (c-di-GMP). The biosynthesis of c-di-GMP by condensation of two GTP molecules is performed by diguanylate cyclases (DGCs). DGCs consist of catalytic GGDEF domains in combination with N-terminal, environment sensing regulatory domains. A significant fraction of DGCs are linked to N-terminal sequences of unknown function indicating that c-di-GMP signaling is linked to numerous undiscovered environmental and cellular signals. In this study structural and biochemical analysis on the DGC YdeH from *E. coli* was undertaken, to elucidate its regulatory mechanism.
Three-dimensional structures of YdeH were determined, which reveal in the regulation of YdeH. The N-terminal sensory domain of YdeH shows a new fold, a four helical bundle, which harbors a zinc-binding site compromised of three histidines and one cysteine. It could be shown that the DGC activity of YdeH is inhibited by zinc binding to the N-terminal sensory domain with an inhibition constant in the femtomolar range. A model for the inhibition of YdeH by zinc is proposed, in which upon zinc binding the linker between the regulatory domain and the enzymatic domain is fixed in a conformation, which prevents the productive encounter of the two GGDEF domains.
In the structures of YdeH, substrate and product binding to the active site could be shown, however the dimeric arrangement of the two DGC domains, each harboring only one half of the active site, are not in a competent constellation. With the help of the determined structures of YdeH a model of a competent dimer was generated, which provides insights into the regulation of YdeH. Product binding to the inhibitory site of YdeH was shown in the crystal structures and inhibition by c-di-GMP was demonstrated in enzymatic experiments. YdeH represents the first example of a biological zinc-sensor that exerts its downstream effects post-transcriptionally and the first example of a metal sensory c-di-GMP signaling protein.

A protocol for the enzymatic large-scale synthesis of c-di-GMP by using the DGC YdeH from *E. coli* was developed and optimized. In contrast to the chemical synthesis of c-di-GMP, enzymatic c-di-GMP production is a one-step reaction that can easily be performed

with the equipment of a standard biochemical lab. The protocol allows the production of milligram amounts of c-di-GMP within one day and paves the way for extensive biochemical and biophysical studies on c-di-GMP-mediated processes.

In biofilms cells are entrapped within a extracellular polymeric matrix. One component of this matrix is the poly-β-1,6-N-Acetyl-glucosamine (poly-1,6-GlcNAc), which is synthesized and exported by the four proteins of the *pgaABCD* operon. PgaC and PgaD are responsible for the synthesis of poly-1,6-GlcNAc and are allosterically regulated by c-di-GMP. The deacetylase PgaB and the outer membrane protein PgaA are involved in the modification and export of the poly-1,6-GlcNAc chain.
For PgaA and PgaB an expression and purification protocol was established and resulted in stable and homogenous proteins. The predicted deacetylase activity of PgaB was demonstrated *in vitro* with an activity assay, which is suitable for rapid screening of different reaction conditions and for the search of inhibitors for PgaB and PgaC, which are of specific pharmaceutical interest.

Abbreviations

ADP	adenosine diphosphate
AEBSF	4-(2-aminoethyl)-benzensulfonylfluorid
ASU	asymmetric unit
ATP	adenosine triphosphate
Bicine	2-(bis(2-hydroxyethyl)amino)acetic acid
Bis-Tris propan	1,3-bis(tris(hydroxymethyl)methylamino)propane
CCD	charge-coupled device
c-di-GMP	bis-(3'- 5')cyclic dimeric guanosine
$C_{12}E_9$	nonaethylene glycol monododecyl ether
Cymal-6	6-cyclohexyl-1-hexyl-β-D-maltoside
CZB	chemoreceptor zinc binding
ddGTP	2',3'-dideoxyguanosine-5'-O-triphosphate
DDM	n-dodecyl-β-D-maltopyranoside
DGC	diganylate cyclase
DM	n-decyl-β-D-maltopyranoside
DNAse	desoxyribonuclease
DsbC-PgaA	DsbC protein fused N-terminally to PgaA
EDTA	ethylenediaminetetraacetic acid
FOM	figure of merit
Fos-12	n-dodecylphosphocholine
FPLC	fast pressure liquid chromatography
GDP	guanosine diphosphate
GlcNAc	N-acetyl-glucosamine
GMP	guanosine monophosphate
GTP	guanosine triphosphate
GTPαS	guanosine-5'-O-(1-thiotriphosphate)
HEPES	hydroxyethylpiperazin-ethansulfonic acid
ICP-MS	inductively coupled plasma mass spectrometry
IPTG	isopropyl-β-D-1-thiogalactopyranoside
I_p-site	primary inhibitory site
I_s-site	secondary inhibitory site

ITC	isothermal titration calorimetry
LDAO	lauryldimethylamine-oxide
MBP	maltose binding protein
MBP-PgaA	maltose binding protein fused N-terminally to PgaA
MES	2-(N-morpholino)ethanesulfonic acid
MPD	2-methyl-2,4 pentanediol
NAD^+	nicotinamide adenine dinucleotide (oxidized form)
NADH	nicotinamide adenine dinucleotide (reduced form)
NCS	non crystallographic symmetry
NMR	nuclear magnetic resonance
OD	optical density
OG	n-octyl-β-D-glucopyranoside
OPOE	n-octylpolyoxyethylene
PAR	4-(2-pyridylazo)resorcinol
PCR	polymerase chain reaction
PDB	protein data bank
PDE	phosphodiesterase
PEG	polyethylene glycol
$PgaA^{TPR}$	tetratrico peptide repeat domain of PgaA
pGpG	5'-phosphoguanylyl-(3'-5')-guanosine
poly-1,6-GlcNAc	poly-β-1,6-N-Acetyl-glucosamine
RMS	root mean square
RT	room temperature
SAD	single wavelength anomalous dispersion
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SeMet	selenomethionine
SLS	Swiss Light Source
SPR	surface plasmon resonance
TCEP	tris-(carboxyethyl)phosphine
TIM	triosephosphate isomerase
TLS	translation libration screw-motion
TPR	tetratrico peptide repeat
Tris	tris(hydroxymethyl)aminomethane
UDP-GlcNAc	uridine diphosphate N-acetylglucosamine
$YdeH^{CZB}$	chemoreceptor zinc binding domain of YdeH
$YdeH^{DGC}$	diguanylate cyclase domain of YdeH

1 Introduction

1.1 The second messenger c-di-GMP

In response to adverse conditions, many bacterial species can switch from a planktonic growth to a surface associated growth mode and form biofilm communities. It is well known that cells in biofilms are protected from physical, chemical or biological stress. Part of this stress tolerance is based on the production of a viscous extracellular matrix, consisting of polysaccharides and pili or fimbriae. The matrix protects biofilms from sheer forces, grazing predators, such as immune cells in host environments and other stressors. In addition, biofilm cells have a remarkable ability to survive antibiotic treatment [1]. Because of this tolerance against antibiotics and the host immune system, biofilm associated infections, such as lung infections in cystic fibrosis patients caused by *Pseudomonas aeruginosa*, or recurring urinary tract infections caused by uropathogenic *Escherichia coli* are notoriously difficult to treat and thus represent a major health problem.
A key factor triggering the formation of biofilms in a multitude of bacterial species, including pathogenic *E. coli* or *P. aeruginosa*, is the second messenger bis-(3'-5')cyclic dimeric guanosine (c-di-GMP) [2]. C-di-GMP was first identified 20 years ago as an activating factor of cellulose synthase in *Acetobacter xylinum* [3], but in the mean time it has been shown to play a central role in the transition between a motile, single-cell state to a sessile, surface attached state found in biofilms [4–6]. Moreover, c-di-GMP controls the virulence of pathogens [2, 7–9], cell cycle progression [10], antibiotic production [11] and other cellular functions.

1.2 The structure of c-di-GMP

C-di-GMP is a two-fold symmetrical molecule that consists of two GMP moieties forming a 12 membered ribose-phosphate ring (Figure 1.1). Several X-ray structures of c-di-GMP have been determined [12–14]. They show similar conformations for the ribose and phosphate moiety in a rather rigid macrocycle. The torsion around the glycosidic bond and therefore the orientation of the guanyl base with respect to the macrocycle is variable. In all the available small molecule crystal structures, c-di-GMP is dimeric with intercalated bases and H-bonding between the guanine N1 and the phosphate group (Figure 1.1). In some

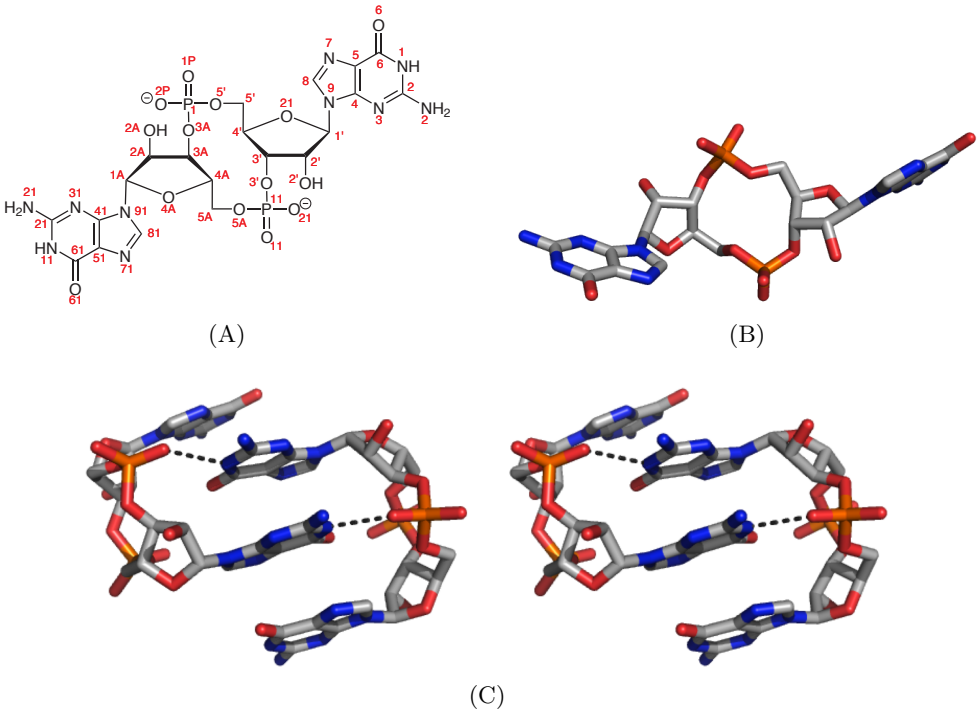


Figure 1.1: Chemical and three-dimensional structure of c-di-GMP. (A) Chemical structure and atom numbering of monomeric c-di-GMP in *anti* conformation. The numbering is adapted from the PDB entry 2von. (B) Three-dimensional structure of monomeric c-di-GMP as it is found in the active site of the EAL phosphodiesterase YkuI from *B. subtilis* (PDB: 2w27). (C) Stereo view of dimeric c-di-GMP as it is bound to the inhibitory site of the diguanylate cyclase PleD from *C. crescentus* (PDB: 1w25). Intermolecular H-bonds are shown as dashed lines.

of the structures a divalent metal ion (Mg^{2+} or Co^{2+}) is coordinated by the N7 atoms of the central bases of the c-di-GMP dimer, which does not change the structure significantly [12, 13]. The same form of dimeric c-di-GMP has been observed in protein complexes, where it binds to the inhibition site of diguanylate cyclases [15–18], PilZ receptors [19, 20] and response regulators [21]. The monomeric form of c-di-GMP was found in c-di-GMP specific phosphodiesterases [22–24], the active site of diguanylate cyclases [15] and in a PilZ receptor [25]. NMR studies in solution at high c-di-GMP concentrations ($> 1\,\mathrm{mM}$) reveal a rich polymorphism, ranging from dimers to several forms of tetramers and octamers in a cation dependent equilibrium [26, 27]. However, a very recent NMR study has shown that c-di-GMP is monomeric at physiological sub-micromolar concentrations [28].

1.3 Biosynthesis and degradation of c-di-GMP

The intracellular concentration of c-di-GMP is controlled by the opposing activities of two signaling enzyme families: diguanylate cyclases (DGCs) and phosphodiesterases (PDEs). DGCs catalyze the condensation of two GTP molecules to c-di-GMP (Figure 1.2). They harbor a characteristic GGDEF domain that is named after a linear motif of amino acids in the active site, which is essential for catalytic activity [29, 30]. The GGDEF domain consists of a five-stranded central β-sheet surrounded by five helices [15]. This fold is related to that of class III nucleotidyl cyclases and type I DNA polymerases, which implies a similar catalytic mechanism using magnesium as metal ions [15, 16]. The active DGC is a dimer of two GGDEF domains, in which both active sites are located at the dimer interface [15, 16]. This allows an antiparallel alignment of two GTP molecules and the synthesis of the two-fold symmetrical c-di-GMP by the formation of two intermolecular phosphodiester bonds. Most of the DGCs show allosteric product inhibition with an inhibition constant in the range of cellular c-di-GMP concentrations [15, 31]. The feedback control avoids excessive GTP consumption and limits the maximal c-di-GMP concentration. The inhibition involves dimeric c-di-GMP binding to a primary inhibition site (I_p) characterized by the RxxD-motif and binding to a secondary inhibition site (I_s) located either on the other GGDEF domain or on a associated regulatory domain [15–17]. Thus c-di-GMP acts as a crosslinker and prevents the encounter of the two active sites.

C-di-GMP is degraded to the linear dinucleotide 5'-phosphoguanylyl-(3'-5')-guanosine (pGpG) by specific PDEs, which can be further degraded to GMP by nonspecific enzymes (Figure 1.2). C-di-GMP specific PDEs either possess an EAL domain or an HD-GYP domain, named after conserved active site residues [32–35]. EAL containing PDEs hydrolyze one ester bond of c-di-GMP to open the c-di-GMP macrocycle [33, 36]. They show high substrate affinity with a K_M in the sub-micromolar range [32, 34, 36] and require Mg^{2+} or Mn^{2+} for catalysis [32, 33]. They fold into a TIM-barrel, in which the active site is located on the bottom of the barrel [22]. The HD-GYP domain proteins are a subfamily of metal-dependent phosphohydrolases and are unrelated to EAL proteins [37]. In contrast to EAL proteins, they catalyze the hydrolysis of phosphodiester bonds to GMP directly [35].

Interestingly, many bacterial genomes encode for dozens of these c-di-GMP signaling proteins, with largely unknown physiological functions. For example, the *Escherichia coli* K-12 genome encodes 29 proteins that harbor a GGDEF, an EAL domain, or both domains, while no HD-GYP domain encoding gene is present. Not all of these proteins are enzymatically active. Some carry alterations of critical active site residues and have adopted alternative functions, that are not necessarily directly related to c-di-GMP signaling [38, 39], but most

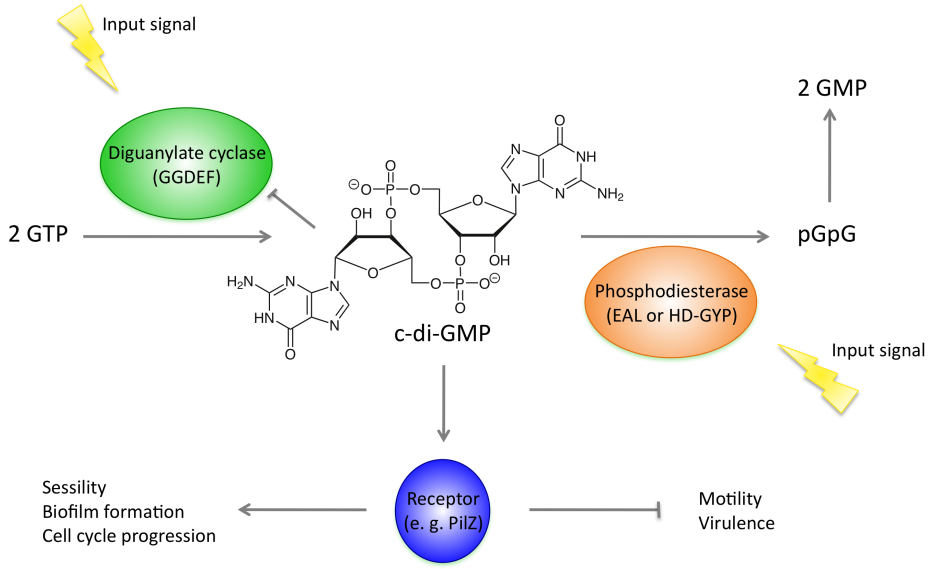


Figure 1.2: C-di-GMP signaling pathways. At the cellular level, c-di-GMP concentration is controlled by diguanylate cyclases (green) and phosphodiesterases (orange). These enzymes respond to internal or external signals that are sensed by the N-terminal accessory domains. C-di-GMP binding to the inhibitory site of diguanylate cyclases results in feedback inhibition. The interaction with different c-di-GMP receptors produces a output signal. Low c-di-GMP concentrations are connected to motility and are required for the expression of virulence genes. High c-di-GMP levels stimulate biofilm formation and are important for cell cycle progression.

GGDEF/EAL domain proteins are predicted, or have been shown, to possess the ability to synthesize or degrade c-di-GMP [40].

1.4 Modulation of the enzymatic activity of c-di-GMP-metabolizing enzymes by signal input domains

The vast majority of c-di-GMP signaling proteins in *E. coli* and other bacterial species harbor N-terminal sensory domains, which are believed to modulate the enzymatic activity of the GGDEF or HD-GYP/EAL output domains after recognizing a specific signal. A significant fraction of DGCs and PDEs contain N-terminal sequences of unknown function indicating that c-di-GMP signaling may be linked to numerous undiscovered environmental and cellular signals. Typical examples of sensory domains are phosphoryl group accepting receiver domains, small molecule binding PAS domains, light sensing BLUF or LOV domains, GAF or HAMP domains, and many others [40]. Most of these domains can also be found in other types of sensory and signaling proteins such as histidine kinases, chemoreceptors of the MCP-type or transcription factors of the one component type [41].
However, the actual signal or ligand that is sensed by these input domains is not known

for almost all c-di-GMP signaling proteins. The few marked exceptions are the PDE DosP, the DGC DosC - both from *E. coli* - and the PDE AxPDE1 from *A. xylinum*, which have been shown to respond to O_2 [42, 43], the DGC AxDGC2 - also from *A. xylinum* - which responds to altered redox conditions via a non-covalently bound FAD cofactor [44], the DGC Lpg1057 from *Legionella pneumophila* which responds to NO via the auxiliary haeme binding protein Hnox1 [45], the GTP sensing PDE CC3396 from *Caulobacter crescentus* [32] and the blue light sensing PDE BlrP1 from *Klebsiella pneumoniae* [23](Figure 1.3). Based on the crystal structures of BlrP1 a activation mechanism was proposed, where light absorption by the flavin molecule cause conformational changes in the BLUF domain, which are transferred to the EAL active site [23]. Other examples are the well studied DGCs PleD and WspR, from *C. crescentus* and *P. aeruginosa*, respectively, which both carry a N-terminal receiver domain and become activated upon phosphorylation of an aspartate residue [30, 46]. Phosphorylation of the REC1 domain of PleD induces structural rearrangement in the REC1-REC2 interfaces, which in turn, allows the tight dimerization of PleD. The dimeric arrangement is a prerequisite for an efficient and productive encounter of the two GTP loaded GGDEF domains to form c-di-GMP [16] (Figure 1.3). However, for PleD and WspR the input cue controlling the activity of their cognate histidine kinases has not yet been identified. For a small number of additional c-di-GMP signaling proteins input signals have been inferred or predicted, but biochemical proof for direct effects are lacking.

1.5 The diguanylate cyclase YdeH from *E. coli*

One example of a c-di-GMP signaling protein for which the input signal has not been identified so far, is the DGC YdeH from *E. coli*. C-di-GMP produced by YdeH posttranslationally controls the production of the polysaccharide adhesin poly-β-1,6-N-Acetyl-glucosamine (poly-β-1,6-GlcNAc) *in vivo* and thereby upregulates *E. coli* biofilm formation [47, 48]. Expression of YdeH is tightly controlled on the mRNA level by the RNA binding protein CsrA, which prevents YdeH translation, and also represses expression of the *pgaABCD* operon, encoding for the poly-β-1,6-GlcNAc biosynthesis machinery [49, 50].
YdeH consists of 296 amino acids with a theoretical mass of 33.9 kDa and contains a C-terminal GGDEF domain. The enzyme is predicted to be active, because it possesses all amino acids, which have been shown to be essential for DGC activity. YdeH has a predicted primary I-site, consisting of a RxxE-motif. This site differs slightly from the canonical RxxD-motif, which is present in the best characterized DGCs PleD from *C. crescentus* and WspR from *P. aeruginosa*. Furthermore the secondary I-site residues, which are involved in the crosslinking of the two GGDEF domains in PleD and WspR, are also present in YdeH. This leads to the conclusion that YdeH might be product inhibited, too.
The N-terminus of YdeH harbors a signal input domain that consists of 126 amino acids and

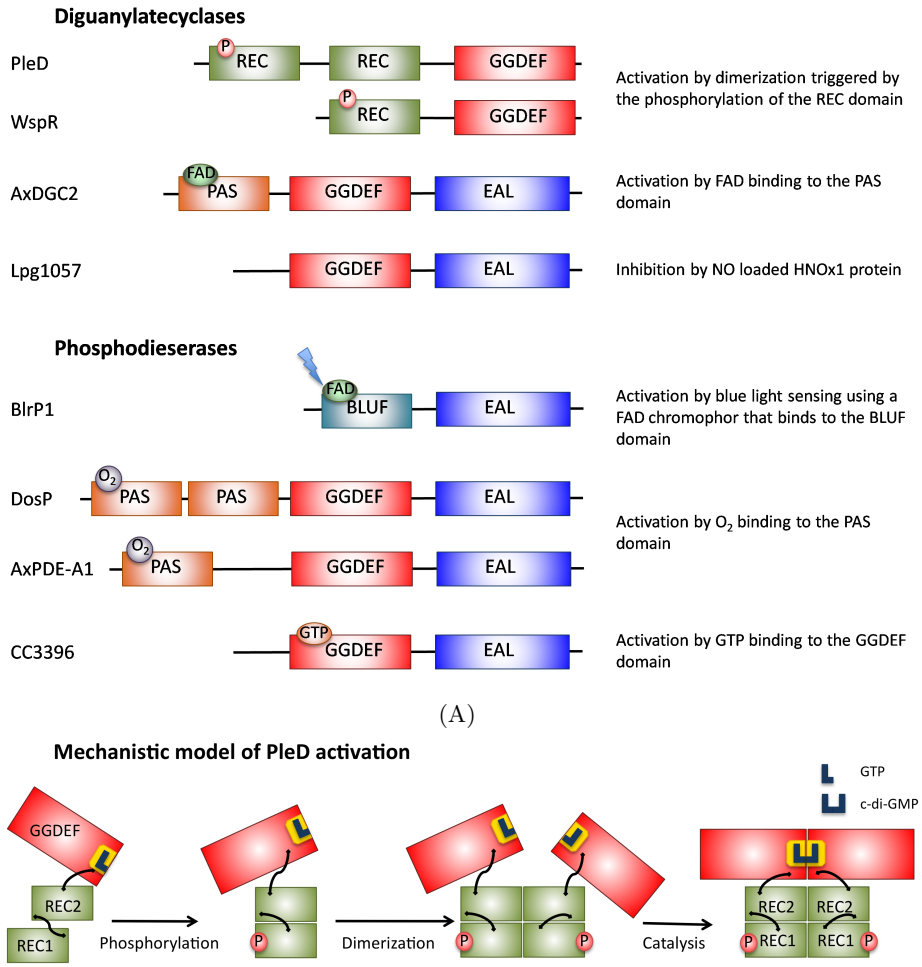


Figure 1.3: Diguanylate cyclases and phosphodiesterases with known input signals. (A) Schematic drawing of the domain organization and the input signals of the DGCs and PDEs. GGDEF and EAL domains are shown in red and blue, respectively. N-terminal regulatory domains are shown in green (REC), orange (PAS) and cyan (BLUF). The regulatory input signals are phosphorylation (P), FAD binding (green), oxygen (purple), GTP (orange) and blue light (blue flash). (B) Mechanistic model of the PleD activation. Phosphorylation of the REC1 domain by a cognate histidine kinase induces structural rearrangement in the REC1-REC2 interface, which in turn, allows tight dimerization of PleD. The dimeric arrangement is a prerequisite for an efficient and productive encounter of the two GTP loaded GGDEF domains to form c-di-GMP. The figure is adapted from [16].

belongs to the Pfam family PB001058 [51]. This family comprises 98 signaling proteins from a variety of bacterial species. 50 % of these proteins are chemosensors that are predicted to be involved in chemotaxis. One such chemosensory protein, TlpD from *Helicobacter pylori*, has recently been shown to bind zinc via the PB001058 domain [52]. Based on the conserved residues involved in zinc binding other PB001058 family members were also predicted to be zinc binding proteins, and the domain was thus named CZB (chemoreceptor zinc binding).

1.6 Zinc and zinc binding proteins

Zinc is a transition metal and serves as a cofactor for many proteins that are involved in a multitude of metabolic processes, including central cellular functions such as transcription and translation [53]. Therefore, zinc represents an essential trace element and all cells possess highly efficient zinc uptake systems. However, due to the reactivity of zinc as a potent Lewis acid, it is essential to strictly prevent nonspecific interaction of freely available zinc with cellular biomolecules. As a consequence, bacteria and other organisms have to keep the intracellular concentration of freely available zinc low, while allowing for full occupation of the highly abundant zinc-cofactor containing proteins. To solve this problem, zinc-binding proteins from bacteria exhibit a remarkably high affinity for the metal, with down to femtomolar dissociation constants, while excess zinc is cleared from the cytoplasm by highly efficient export systems [54]. The activity of zinc import and export systems is sensitively balanced to ensure the maintenance of tight zinc homeostasis over a wide range of ambient zinc concentrations that can vary between low nanomolar and high micromolar levels [55].

In the majority of zinc-binding proteins, the metal either functions as a catalyst or as a pure structural factor, necessary for the formation of the proper secondary, tertiary or quaternary structure of the zinc-binding protein [56]. The prototypical chemical function of a catalytic zinc is to activate water molecules and make them available for a large variety of biochemical reactions. The coordination number is four or five with a tetrahedral or trigonal-bipyramidal geometry. In every case, water is present as a ligand. In these zinc enzymes, the charges of the coordinating amino acids modulate the Lewis acidity of the zinc ion, which can activate the bound water molecule. Structural zinc ions have four protein ligands and no bound water molecule arranged in a tetrahedral manner [56]. Typical zinc-liganding residues are histidine, cysteine, glutamate and aspartate.
Identifying a site as structural or catalytic is not as straightforward as it appears. Zinc sites that have been classified as structural sites based on the presence of four protein ligands, were shown to have additional functions such as redox activity [57]. Furthermore, the metal in zinc-responsive signal transduction proteins has neither a catalytic nor a pure structural function. Instead, the zinc allosterically controls the activity of an effector domain, for

example the DNA-binding domain of a transcription factor, or the histidine kinase activity of two-component-systems signaling proteins [55, 58]. These regulators are - among other functions - partially responsible for the maintenance of zinc homeostasis by modulating the abundance of zinc exporters and importers in response to changes of the ambient zinc concentrations [55]. Two zinc-responsive transcription factors are known in the bacterial kingdom and both have representatives in *E. coli*: the MerR-type transcription activator ZntR, which upon binding of zinc induces expression of the zinc exporter coding gene *zntA* [59], and the Fur-like repressor Zur, which represses transcription of several genes coding for zinc uptake systems and other genes facilitating growth under zinc-limited conditions [60, 61]. While ZntR binds two zinc ions via a single binding site that is composed of residues from both protomers in the dimeric protein [62], a close homologue of Zur, FurB from *Mycobacterium tuberculosis*, was found to bind three zinc ions via three distinct and prototypical tetrahedral sites per monomer [63]. The binding constant of ZntR was determined to be in the femtomolar range [64]. Since the binding affinity of zinc in zinc proteins is very high, kinetic mechanisms must exist for the dissociation of the tightly bound zinc [57].

1.7 C-di-GMP receptors

After DGCs have synthesized c-di-GMP in respond to external or internal signals, these input signals have to be further transferred via c-di-GMP binding to an effector component by inducing a structural or functional change. As c-di-GMP controls a wide range of cellular functions and processes, c-di-GMP receptors are highly diverse. Only a few types of different c-di-GMP receptors are currently known. The best-studied class is the PilZ proteins, which seem to be activated by c-di-GMP and to function by protein-protein interaction. Structural studies revealed that c-di-GMP binding to the PilZ protein induces dramatic conformational changes, which makes the PilZ domain more compact and provides a novel surface for protein-protein interactions [25, 65, 66]. PilZ proteins can be standalone proteins or can be linked to other domains, that generate a molecular output. PilZ proteins are involved in regulation of biofilm formation, cellulose biosynthesis, motility, extracellular enzyme production and virulence [65–68].

Another type of receptor proteins is the transcription factor FleQ of *P. aeruginosa*, which is repressed by c-di-GMP [69], and PelD of *P. aeruginosa*, which is activated by binding of c-di-GMP to a site similar to the I-site of DGCs [70].

Degenerated GGDEF proteins can also act as c-di-GMP receptors, where c-di-GMP binding to the I-site alters protein function. PopA from *C. crescentus*, recruits the cell cycle regulator CtrA to the cell pole after c-di-GMP binding to the I-site, thereby targeting CtrA for

degradation [10]. Another I-site dependent c-di-GMP receptor is CdgG from *Vibrio cholera*, that controls biofilm formation and motility [71].

An additional class of c-di-GMP receptors are riboswitches. The conserved GEMM domain is present in the 5'-untranslated regions of different mRNAs, which regulate gene expression via c-di-GMP binding [72].

1.8 A c-di-GMP responsive system: the *pgaABCD* operon of *E.coli*

In biofilms, cells are entrapped within an extracellular polymeric matrix [73, 74]. While this matrix contains a variety of components affecting its properties, the structural integrity of biofilm often depends on polysaccharides [48, 75–77]. Poly-β-1,6-N-Acetyl-glucosamine (poly-1,6-GlcNAc) is a homopolymer originally found in *Staphylococcus epidermidis* [78], but it can also be isolated from *E. coli* [48]. Poly-1,6-GlcNAc is an important cell-to-cell and cell-to-surface adhesion molecule in biofilms and is essential for the formation of the cellular architecture of *E. coli* biofilm microstructure [79, 80]. Furthermore, it is crucial during the initial stages of biofilm development and has effects on host-microbe interactions [81–84].
The production of poly-1,6-GlcNAc in *E. coli* is dependent on the *pgaABCD* operon [48] and is regulated by CsrA and NahR. The level of poly-1,6-GlcNAc is posttranscriptionally repressed by the carbon storage regulator CsrA that binds to the coding region of *pgaA* mRNA [50, 85]. Poly-1,6-GlcNAc synthesis requires the DNA-binding protein NahR, which activates the *pgaABCD* transcription in response to high pH and high Na^+ concentrations [86].

The *pgaABCD* operon encodes for four proteins, that are responsible for the synthesis and export of poly-1,6-GlcNAc. PgaC is a glycosyltransferase and has five predicted transmembrane helices anchoring the protein in the inner membrane and exposing a soluble domain to the cytoplasm. Glycosyltransferases are involved in the synthesis of sugar chains and in the biosynthesis of glycoproteins and glycolipids. They transfer a monosaccharide from an activated sugar donor, for example UDP-GlcNAc, to a saccharide, a protein or a lipid. Generally glycosyltransferases fold into a eight stranded β-sheet flanked by α-helices. For catalysis, a metal ion is required [87]. *In vivo* experiments have shown, that PgaC is essential for poly-1,6-GlcNAc accumulation [88].
PgaD is a small 137 amino acid containing inner membrane protein with two predicted transmembrane helices at its N-terminus. For PgaD, no structural related proteins could

be found. Like PgaC, PgaD is required for the poly-1,6-GlcNAc synthesis [88]. The homologue of PgaD, HmsS in *Yersinia pestis*, is no longer able to produce poly-1,6-GlcNAc, if conserved residue in the predicted transmembrane helix 2 are mutated [89]. Thus, this membrane spanning region might be a putative binding site for PgaC to promote the glycosyltransferase activity. Very recent studies have shown, that PgaD interacts with PgaC in a c-di-GMP dependent manner and that this interaction is required for the synthesis of poly-1,6-GlcNAc (S. Steiner, unpublished data).
The 672 amino acid containing PgaB has a 28 amino acid signal sequence containing a lipo box, which anchors the protein in the outer membrane by a lipid anchor. PgaB has an N-terminal deacetylation domain and a C-terminal domain of unknown function. Polysaccharide deacetylases cleave acetyl groups from sugars and thereby modify cell surface properties and play a role in the protection against host defences [90–92]. Deacetylases fold in a TIM-barrel with the active site at the C-terminal ends of the β-sheet [93]. Most of them require a metal ion for catalysis [94, 95]. PgaB needs both domains for the export of the poly-1,6-GlcNAc chains to the cell surface [88]. Because the active site residues of the deacetylation domain are essential for the export of the poly-1,6-GlcNAc, deacetylation is required for the transport through the outer membrane [88]. The deacetylation level was measured either with NMR resulting in 3 % deacetylated sugar groups [48] or with a ninhydrin assay which gave a deacetylation level of 22 % [88].
The last member of the *pgaABCD* operon is the outer membrane protein PgaA. It contains 807 amino acids including a 33 amino acid long signal sequence for the transport to the outer membrane. In addition to the C-terminal membrane domain, that is predicted to form a 16-stranded β-barrel, PgaA contains a large N-terminal soluble tetratrico peptide repeat (TPR) domain. The TPR domain consists of a 34 amino acid motif, which folds into two antiparallel α-helices linked by a turn. This motif occurs in tandem of three to 16 copies, where adjacent repeats stack in a parallel fashion. This domain is found in a variety of proteins, where it is involved in protein-protein interactions. *In vivo* studies have shown, that PgaA is essential for the export of poly-1,6-GlcNAc [88].

In Figure 1.4, a model of the proteins arrangement from the *pgaABCD* operon is shown. The inner membrane proteins PgaC and PgaD are essential for the synthesis of the sugar chain and these proteins interact probably via their transmembrane helices. So far, it is not clear how the sugar chain that is synthesized in the cytosol crosses the inner membrane. The poly-1,6-GlcNAc might be passed through a pore formed by the transmembrane helices of PgaC and PgaD. In the periplasm, poly-1,6-GlcNAc is partially deacetylated by PgaB. As PgaB and PgaA are needed for the export of the sugar chain, it is likely that PgaB binds to the TPR domain of PgaA, which is known to mediate protein-protein interactions in other TPR-domain containing proteins.

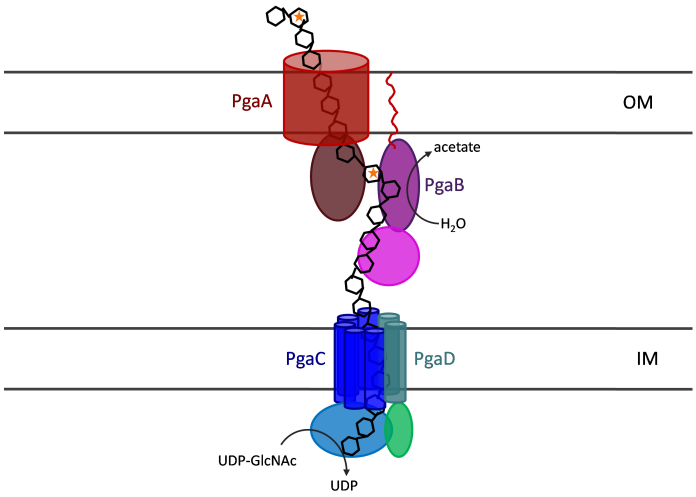


Figure 1.4: Model of the synthesis and export of poly-1,6-GlcNAc by the proteins from the *pgaABCD* operon. The orange asterisks mark deacetylated sugar groups. (OM: outer membrane, IM: inner membrane).

1.9 Chemical and enzymatic synthesis of c-di-GMP

In order to carry out extensive studies on c-di-GMP mediated processes, sufficient supply of this compound is crucial. The chemical synthesis of c-di-GMP is a multistep reaction process, which makes it time consuming and expensive [3, 26, 96–100]. Most of them make use of phosphotriester, phosphoamidite or H-phosphonate chemistry, which is based on air or water sensitive reagents. The purification of synthetic c-di-GMP involves several chromatographic steps and results in very low product yield.
In contrast, enzyme catalyzed synthesis of c-di-GMP from GTP by DGCs is straightforward and efficient. *In vitro* enzymatic production of c-di-GMP has been described by the use of the DGCs PleD, VCA0956 and WspR [32, 34, 69, 101, 102]. However, all these DGCs show allosteric inhibition at micromolar product concentration and poor stability after purification, which do not qualify them for the production of large amounts of the second messenger. Recently, a thermophilic DGC has been described for the enzymatic synthesis of c-di-GMP. In order to improve thermostability, only the DGC domain of the protein was used, and product inhibition was avoided by site directed mutagenesis of the allosteric site [103].

1.10 Aims of this work

Structural and enzymatic characterizations of diguanylate cyclase have already been initiated earlier and provided insights in the catalysis and in the product inhibition of DGCs. Major work was done on the DGCs PleD from *C. crescentus* and WspR from *P. aeruginosa*, which both contain a REC domain as the N-terminal sensory domain. These DGCs are regulated by phosphorylation of the REC domain followed by dimerization to receive active protein. Several DGCs harbor undefined N-terminal domains and it is poorly understood, wether these domains are able to form permanent dimers or wether dimer formation is regulated by input signals. A few crystal structures of DGCs are available and illuminate substrate and product binding in the active site. However, a structure of a catalytically competent DGC dimer is still missing. Structural and enzymatic studies showed, that most DGCs are allosterically product inhibited by the crosslinking of domains in an unproductive state. In the I_p-site a conserved RxxD-motif is involved. However it is unclear, if DGCs lacking this motif are also product inhibited.
In the present thesis, the DGC YdeH from *E.coli* was used to gain more information in general by applying structural and biochemical approaches. The main focus was on crystallizing YdeH to obtain structural information of the competent state and on the N-terminal sensory domain. All experiments were carried out to elucidate the following questions: Is YdeH an active diguanylate cyclase *in vitro*? Is it a permanent dimer? Does it show product inhibition, although it has a slightly different primary inhibition motif? What is the input signal for YdeH and how is YdeH regulated?

Furthermore, a protocol for the enzymatic production and purification of c-di-GMP should be developed to generate large amounts of c-di-GMP, that is needed to carry out *in vitro* research on c-di-GMP related topics.

The second part of the thesis deals with PgaA and PgaB, two proteins from the *pgaABCD* operon, which is involved in c-di-GMP dependent exopolysaccharide production. A protocol for the expression and purification of these two proteins should be developed to carry out structural and biochemical investigations. The questions addressed in this part of the present thesis were the following: Can a protein-protein interaction between PgaA and PgaB be demonstrated to prove the model for the export of the sugar chain? Is PgaB an active deacetylase? The main objective was to obtain the crystal structures of PgaA and PgaB. It is likely that the crystal structure of PgaA will help to get more precise information about the export mechanism of the sugar. Furthermore, the structure of PgaB will analyze the deacetylation process in detail and will probably provide more insights into the function of its so far unidentified C-terminal domain. Furthermore both proteins are promising drug targets and therefore structural information of both proteins is of specific pharmaceutical interest.

2 Methods

2.1 Cloning

Cloning of individual domains of YdeH

Coding regions corresponding to YdeH^{CZB} (residues 1-126) and YdeH^{GGDEF} (residues 127-296) from *E. coli* were amplified by standard polymerase chain reaction (PCR) using the YdeH containing plasmid pET28a/YdeH produced by Dr. A. Böhm as template DNA [47]. The DNA inserts were cloned into the pET28a expression plasmid (Novagen), yielding C-terminally hexahistidine-tagged proteins.

Mutagenesis

YdeH mutants were generated using the QuikChange II Site-directed Mutagenesis Kit (Agilent Technologies) based on the pET28a/YdeH plasmid according to manufacturers instructions.

Cloning of PgaA

The coding region corresponding to PgaA from *E.coli* with and without the signal sequence as well as the coding region corresponding to the soluble TPR domain of PgaA (residues 33 - 519) were amplified by colony PCR using the *E. coli* strain MG1655 as template. DNA inserts containing full-length PgaA including the signal sequence were cloned in the expression plasmids pET28a (Novagen) and pColdIV (Takara) yielding C-terminally hexahistidine-tagged protein. DNA inserts of PgaA without the signal sequence were cloned into the expression plasmids pET22b (Novagen) generating C-terminally hexahistidine-tagged protein, into pMal-p5X plasmid (NEB) yielding N-terminally MBP fusion protein and in the pET40b plasmid (Novagen) resulting in N-terminally hexahistidine-tagged DsbC fusion protein. The MBP moiety and the hexahistidine-tagged DsbC moiety were cleavable by Factor Xa or thrombin, respectively.

Cloning of PgaB

The coding region corresponding to PgaB from *E.coli* without the lipid anchor (residues 22-651) was amplified by colony PCR using the *E. coli* strain MG1655 as template. DNA

inserts were cloned in the expression plasmids pET21b (Novagen) and pET28a (Novagen) yielding N-terminally and C-terminally hexahistidine-tagged protein, respectively.

2.2 Protein production

2.2.1 Protein expression

Expression tests of YdeH

The pET28a/YdeH plasmid was transformed in different *E.coli* strains BL21(DE3)pLysS (Novagen), BL21(DE3)pGroEL, Rosetta(DE3) (Novagen) and ArcticExpress(DE3) (Agilent Technologies). 50 mL LB-medium supplemented with antibiotics was inoculated with overnight culture to a starting OD_{600} of 0.1. The cells were grown at 37 °C. At an OD_{600} of 0.7, the protein expression was induced with 1 mM IPTG (isopropyl-β-D-1-thiogalactopyranoside). For the expression test at lower temperatures, cells were grown at 37 °C until an OD_{600} of 0.5 was reached before temperature was reduced to 18 °C. Before and at several time points after induction samples were taken, which were centrifuged (10 000 g, 10 min, room temperature) and the cell pellet was resuspended in YdeH-Ni-A-buffer (50 mM NaH_2PO_4, pH 7.5, 200 mM NaCl, 10 mM imidazole, 50 mM L-glutamic acid, 50 mM L-arginine), where the volume of the used buffer corresponds to the cell density of the sample. The cells were lysed by sonication and centrifuged (10 000 g, 10 min, RT). The supernatant containing the soluble protein fraction, was loaded on an SDS-PAGE to follow YdeH expression.

Expression of YdeH

YdeH was produced using the *E.coli* Rosetta(DE3) strain transformed with the pET28a/YdeH plasmid. Cells were cultivated at 37 °C in LB-medium supplemented with kanamycin (100 mg/mL) and chloramphenicol (34 mg/mL) until the OD_{600} had reached 0.7. Gene expression was induced by adding 1 mM IPTG. After 4 h of incubation, the cells were harvested by centrifugation (6800 g, 10 min, 4 °C) and the pellets were frozen at −20 °C.

Expression of individual domains of YdeH

Individual domains of YdeH were expressed in the *E.coli* Rosetta(DE3) strain transformed with the plasmids pET28a/YdeH^{CZB} or pET28a/YdeH^{GGDEF}. Cells were grown in LB-medium supplemented with kanamycin (100 mg/mL) and chloramphenicol (34 mg/mL) at 37 °C until the OD_{600} had reached 0.5. The temperature was then reduced to 18 °C and gene expression was induced by adding 1 mM IPTG. After 20 hours, cells were harvested by centrifugation (6800 g, 10 min, 4 °C) and the pellets were frozen at −20 °C for further use.

Expression of selenomethionine labeled YdeH^{CZB}

Selenomethionine substituted YdeH^{CZB} was expressed in the auxotrophic strain B834(DE3) (Novagen) in LeMaster-medium [104]. Overnight culture, grown in LB-medium supplemented with kanamycin (100 mg/mL) and chloramphenicol (34 mg/mL) was centrifuged (4200 g, 10 min, 4 °C) and resuspended in LeMaster-medium. After a second centrifugation step, the cells were again resuspended in LeMaster-medium and used for inoculation of the main culture (LeMaster-medium supplemented with kanamycin (100 mg/mL) and chloramphenicol (34 mg/mL)). The bacterial culture was incubated at 37 °C until the OD_{600} had reached 0.6, after which the temperature was reduced to 18 °C and gene expression was induced by adding 1 mM IPTG. Incubation was continued for 20 h, before cells were harvested by centrifugation (6800 g, 10 min, 4 °C) and frozen at −20 °C for further use.

Expression tests of PgaA

PgaA containing expression plasmids were transformed in different *E.coli* strains. 125 mL LB-medium supplemented with antibiotics was inoculated with 1 mL overnight culture. Cells were grown at 37 °C. At an OD_{600} of 0.7, protein expression was induced with 1 mM IPTG. For the expression test at lower temperatures, cells were grown at 37 °C until an OD_{600} of 0.5 was reached before the temperature was reduced to 18 °C. Before and at several time points after induction 25 mL samples were taken, which were centrifuged (4500 g, 10 min, 4 °C) and the cell pellets were resuspended in 8 mL extraction-buffer (50 mM NaH_2PO_4, pH 7.5, 200 mM NaCl). Cells were lysed with a French press at 15 000 psi and centrifuged (4500 g, 15 min, 4 °C) to remove unbroken cells and inclusion bodies. The supernatant was ultracentrifuged (100 000 g, 30 min, 4 °C) to collect the total membrane fraction that was resuspended in 1 mL extraction-buffer. After solubilization of the inner membrane fraction by addition of 1 % N-lauroyl sarcosine, the outer membrane fraction was pelleted by ultracentrifugation (107 000 g, 20 min, 4 °C) and resuspended in 1 mL extraction-buffer. PgaA was extracted from the outer membrane with 0.2 % lauryldimethylamine-oxide (LDAO). An additional centrifugation step (107 000 g, 20 min, 4 °C) was performed to remove the residual membrane. 15 μL supernatant, containing PgaA extracted from the outer membrane, were loaded on an SDS-PAGE to follow the expression.

Expression of PgaA

PgaA was produced using the *E.coli* BL21(DE3) strain transformed with the pET28a/PgaA plasmid. Cells were cultivated at 37 °C in LB-medium supplemented with kanamycin (100 mg/mL) until the OD_{600} had reached 0.7. Protein expression was induced by adding 1 mM IPTG. After 3 h of incubation, cells were harvested by centrifugation (6800 g, 10 min, 4 °C) and the pellets were frozen at −20 °C.

DsbC-PgaA fusion protein was expressed in BL21(DE3) cells containing the pET40b/PgaA plasmid. Cells were grown at 37 °C in LB-medium supplemented with kanamycin (100 mg/mL) and chloramphenicol (34 mg/mL) until the OD_{600} had reached 0.5. Temperature was reduced to 18 °C and protein expression was induced with 0.1-1 mM IPTG. After 20 h of incubation, cells were harvested by centrifugation (6800 g, 10 min, 4 °C) and pellets were frozen at −20 °C.

Expression tests of PgaA^{TPR}

The pET28a/PgaA^{TPR} plasmid was transformed in different *E.coli* strains BL21(DE3) (Novagen), BL21(DE3)pLysS (Novagen), Rosetta(DE3) (Novagen), C43(DE3) (Lucigen) and ArcticExpress(DE3)(Agilent Technologies). 50 mL LB-medium supplemented with antibiotics were inoculated with 1 mL overnight culture. Cells were grown at 37 °C. At an OD_{600} of 0.7, protein expression was induced with 1 mM IPTG. For the expression test at lower temperatures, cells were grown at 37 °C until an OD_{600} of 0.5 was reached before temperature was reduced to 18 °C. Before and at several time points after induction, samples were taken, which were centrifuged (10 000 g, 10 min, RT) and the cell pellet was resuspended in 8 mL 50 mM NaH_2PO_4, pH 7.5, 200 mM NaCl. Cells were lysed with a French press at 15 000 psi and centrifuged (4500 g, 15 min, RT) to remove unbroken cells and inclusion bodies. The supernatant, containing the soluble protein fraction, was loaded on an SDS-PAGE to follow PgaA^{TPR} expression.

Expression of PgaA^{TPR}

PgaA^{TPR} was produced using the *E.coli* Rosetta(DE3) strain transformed with the pET28a/PgaA^{TPR} plasmid. Cells were cultivated at 37 °C in LB-medium supplemented with kanamycin (100 mg/mL) and chloramphenicol (34 mg/mL) until the OD_{600} had reached 0.5. Temperature was reduced to 18 °C and protein expression was induced with 1 mM IPTG. After 20 h of incubation, cells were harvested by centrifugation (6800 g, 10 min, 4 °C) and the pellets were frozen at −20 °C.

Expression tests of PgaB

pET21b/PgaB and pET28a/PgaB were transformed in different *E. coli* strains BL21(DE3) (Novagen), BL21(DE3)pLysS (Novagen), Rosetta(DE3) (Novagen), C43(DE3) (Lucigen) and ArcticExpress(DE3)(Agilent Technologies). 125 mL LB-medium supplemented with antibiotics were inoculated with 1 mL overnight culture. Cells were grown at 37 °C. At an OD_{600} of 0.7 protein expression was induced with 1 mM IPTG. For the expression test at lower temperatures, cells were grown at 37 °C until an OD_{600} of 0.5 was reached before temperature was reduced to 18 °C. Before and at several time points after induction, 25 mL

samples were taken, which were centrifuged (10 000 g, 10 min, RT) and the cell pellet was resuspended in 8 mL 50 mM NaH_2PO_4, pH 7.5, 200 mM NaCl. Cells were lysed with a French press at 15 000 psi and centrifuged (4500 g, 15 min, RT) to remove unbroken cells and inclusion bodies. The supernatant, containing the soluble protein fraction, was loaded on an SDS-PAGE to follow PgaB expression.

Expression of PgaB

PgaB was produced using the *E.coli* BL21(DE3) strain transformed with the pET21b/PgaB plasmid or using the *E.coli* Rosetta(DE3) strain transformed with the pET28a/PgaB plasmid. Cells were cultivated at 37 °C in LB-medium supplemented with ampicillin (100 mg/mL) or with kanamycin (100 mg/mL) and chloramphenicol (34 mg/mL), respectively until the OD_{600} had reached 0.5. Temperature was reduced to 18 °C and protein expression was induced with 1 mM IPTG. After 20 h of incubation, cells were harvested by centrifugation (6800 g, 10 min, 4 °C) and the pellets were frozen at −20 °C.

2.2.2 Protein purification

Purification of YdeH

Prior to lysis, the frozen cell pellets were thawed and resuspended in YdeH-Ni-A buffer (50 mM NaH_2PO_4, pH 7.5, 200 mM NaCl, 10 mM imidazole, 50 mM L-glutamic acid, 50 mM L-arginine). After addition of DNAse (2.5 mg/mL) and EDTA-free protease-inhibitor cocktail (Roche, 1 tablet/50 mL buffer), cells were disrupted with a French press at 15 000 psi. The lysate was cleared by centrifugation (28 000 g, 45 min, 4 °C), and the supernatant was filtered (0.22 μm) and loaded onto a 5 mL HisTrap column (GE Healthcare). After washing the column with YdeH-Ni-A buffer (10 column volumes (CV)), the protein was eluted with a linear gradient of imidazole from 10 to 500 mM in 10 CV. YdeH containing fractions were pooled and concentrated to 1 mL and further purified by size-exclusion chromatography using a HiLoad-16/60-Superdex-75 prep grade column (GE Healthcare) and YdeH-GF buffer (20 mM Tris-HCl, pH 7.6, 150 mM NaCl, 50 mM L-glutamic acid, 50 mM L-arginine). The purification process was monitored by SDS-PAGE. Protein concentration was determined by measuring the absorbance at 280 nm ($\epsilon_{280} = 39\,880\,M^{-1}cm^{-1}$).

Purification of individual domains of YdeH

For the purification of individual domains of YdeH ($YdeH^{CZB}$ and $YdeH^{GGDEF}$) the same procedure as described above for full-length YdeH was used.

Table 2.1: Detergents and concentrations used for the stepwise extraction of PgaA from the outer membrane.

Detergent	Concentration %			
	Step 1	Step 2	Step 3	Step 4
DM	0.15	0.3	0.6	1
DDM	0.01	0.02	0.03	0.04
LDAO	0.05	0.1	0.2	0.3
OG	0.5	1	2	3
OPOE	0.5	1	2	3
$C_{12}E_9$	0.005	0.01	0.02	0.04
FOS-12	0.1	0.2	0.3	0.4
Cymal-6	0.04	0.08	0.15	0.3

Purification of selenomethionine labeled YdeH^{CZB}

Purification of selenomethionine labeled YdeH^{CZB} was done with the same protocol as for wildtype YdeH^{CZB}, but with 0.5 mM Tris-(carboxyethyl)phosphine (TCEP) added to all buffers.

Extraction test of PgaA

Cells were resuspended in 8 mL extraction-buffer (50 mM NaH_2PO_4, pH 7.5, 200 mM NaCl), lysed with a French press at 15 000 psi and centrifuged (4500 g, 15 min, 4 °C) to remove unbroken cells and inclusion bodies. The supernatant was ultracentrifuged (100 000 g, 30 min, 4 °C) for collection of the total membrane fraction, that was resuspended in 1 mL extraction-buffer. After solubilization of the inner membrane fraction by addition of 1 % N-lauroyl sarcosine, the outer membrane fraction was pelleted by ultracentrifugation (107 000 g, 20 min, 4 °C). For the stepwise extraction of PgaA from the outer membrane different detergents using increasing concentrations were tested, which are listed in Table 2.1. Therefore, the outer membrane was resuspended in 1 mL extraction buffer supplemented with the corresponding detergent. After centrifugation (107 000 g, 20 min, 4 °C), the pellet was again resuspended in 1 mL extraction buffer containing a higher detergent concentration (see Table 2.1). This extraction step was repeated two more times. The supernatant of each extraction step was loaded on an SDS-PAGE to follow the extraction of PgaA from the outer membrane.
To find the optimal buffer composition for the extraction of PgaA from the outer membrane, the extraction was performed in one step with 0.5 % LDAO in the following different extraction buffers: 100 mM MES, pH 6.5, 100 mM Na_2HPO_4, pH 7.5, 100 mM Tris-HCl, pH 8.0 or 100 mM Bicine, pH 9.0 with and without 200 mM NaCl.

Purification of PgaA

Prior to lysis, the frozen cell pellets from 5 L culture were thawed and resuspended in 50 mL extraction buffer (100 mM NaH_2PO_4, pH 8.0, 200 mM NaCl). After addition of DNAse (2.5 mg/mL), cells were disrupted with a French press at 15 000 psi. Two centrifugation steps (4500 g, 15 min, 4 °C followed by 100 000 g, 30 min, 4 °C) allowed the elimination of unbroken cells and inclusion bodies and subsequent collection of the total membrane fraction that was resuspended in 40 mL extraction buffer. After solubilization of the inner membrane fraction by addition of 1 % N-lauroyl sarcosine, the outer membrane fraction was pelleted by ultracentrifugation (107 000 g, 20 min, 4 °C) and resuspended in 20 mL extraction buffer. PgaA membrane extraction was achieved by stepwise solubilization of the outer membrane using increasing detergent concentrations (from 0.05 % to 0.5 %) of LDAO. Solubilized PgaA was filtered (0.22 μm) and loaded onto a 5 mL HisTrap column (GE Healthcare) in a circle for 18 h. After washing the column with extraction buffer containing 0.1 % LDAO (10 CV), protein was eluted with a linear gradient of imidazole from 0 to 500 mM in 10 CV. PgaA containing fractions were pooled and concentrated to 0.5 mL and further purified by size-exclusion chromatography using a Superdex-200 10/30 column (GE Healthcare) and PgaA-GF buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.1 % LDAO). The purification process was monitored by SDS-PAGE. Protein concentration was determined by measuring the absorbance at 280 nm ($\epsilon_{280} = 200\,420\,M^{-1}cm^{-1}$).

Purification of PgaA as DsbC fusion

For the purification of DsbC-PgaA the same procedure as for PgaA was used as described above.

Purification of $PgaA^{TPR}$

Prior to lysis, the frozen cell pellets were thawed and resuspended in $PgaA^{TPR}$-Ni-A buffer (50 mM NaH_2PO_4, pH 7.5, 200 mM NaCl, 10 mM imidazole, 50 mM L-glutamic acid, 50 mM L-arginine). After addition of DNAse (2.5 mg/mL), cells were disrupted with a French press at 15 000 psi. The lysate was cleared by centrifugation (28 000 g, 45 min, 4 °C), and the supernatant was filtered (0.22 μm) and loaded onto a 5 mL HisTrap column (GE Healthcare). After washing the column with $PgaA^{TPR}$-Ni-A buffer (10 CV), protein was eluted with a linear gradient of imidazole from 10 to 500 mM in 10 CV. $PgaA^{TPR}$ containing fractions were pooled and concentrated to 1 mL and further purified by size-exclusion chromatography using a HiLoad-26/60-Superdex-200 prep grade column (GE Healthcare) and $PgaA^{TPR}$-GF buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 50 mM L-glutamic acid, 50 mM L-arginine). The purification process was monitored by SDS-PAGE. Protein concentration was determined by measuring the absorbance at 280 nm ($\epsilon_{280} = 99\,000\,M^{-1}cm^{-1}$).

Purification of PgaB

Prior to lysis, the frozen cell pellets were thawed and resuspended in PgaB-Ni-A buffer (50 mM NaH_2PO_4, pH 7.5, 200 mM NaCl, 10 mM imidazole). After addition of DNAse (2.5 mg/mL), cells were disrupted with a French press at 15 000 psi. The lysate was cleared by centrifugation (28 000 g, 45 min, 4 °C), and the supernatant was filtered (0.22 μm) and loaded onto a 5 mL HisTrap column (GE Healthcare). After washing the column with PgaB-Ni-A buffer (10 CV), protein was eluted with a linear gradient of imidazole from 10 to 500 mM in 10 CV. PgaB containing fractions were pooled and concentrated to 3 mL and further purified by size-exclusion chromatography using a HiLoad-26/60-Superdex-200 prep grade column (GE Healthcare) and PgaB-GF buffer (50 mM HEPES, pH 7.5, 50 mM NaCl). The purification process was monitored by SDS-PAGE. Protein concentration was determined by measuring the absorbance at 280 nm ($\epsilon_{280} = 47\,220\,M^{-1}cm^{-1}$).

2.3 Protein biochemistry

2.3.1 Zinc removal from YdeH

To obtain zinc-free protein, YdeH (1 mg/mL) was incubated with 50 mM ethylenediaminetetraacetic acid (EDTA) for 16 h at 4 °C. To remove the chelator, a gel filtration was performed using a HiLoad-16/60-Superdex-75 prep grade column (GE Healthcare) with 20 mM Tris-HCl, pH 7.6, 150 mM NaCl, 50 mM L-glutamic acid and 50 mM L-arginine as running buffer.

2.3.2 PAR assay

To control the release of zinc from YdeH after treatment with EDTA, the amount of zinc bound to the protein was measured with a PAR assay [105]. Therefore 5 μM YdeH was incubated in 50 mM Tris-HCl, pH 8.0 with 6 M guandinium hydrochloride and 0.5 mM methylmercury(II)acetate for 16 h. After addition of 100 μM 4-(2-pyridylazo)resorcinol (PAR), the absorption at 495 nm was measured. The zinc content was quantified with a $ZnCl_2$ standard.
To measure the amount of zinc in solution, a variation of the PAR assay was used. The procedure was the same as described above, but the addition of guandinium hydrochloride and methylmercury(II)acetate was omitted.

2.3.3 DGC assay

DGC activity measurement using ion-exchange chromatography

To test for diguanylate activity, 2 μM purified YdeH was incubated with 100 μM GTP in 50 mM Tris-HCl, pH 8.0, 50 mM NaCl and 5 mM $MgCl_2$. The reaction was stopped after 2 or 15 min by heating the sample to 99 °C for 2 min. Subsequently, 100 μL of the reaction mixture were diluted in 900 μL 5 mM NH_4HCO_3, pH 8.0, filtered (0.22 μm) and loaded on an ion-exchange column (ResourceQ 1 μL, GE Healthcare). The nucleotides were separated with a gradient from 0.005 to 1 M NH_4HCO_3, pH 8.0, in 14 CV. The elution of the reaction was compared to the elution profiles of GTP (Sigma) and c-di-GMP (Biolog, Bremen).
The reaction product of YdeH was further analyzed by high-performance liquid chromatography coupled to mass spectrometry. For this, the samples were diluted with 10 mM NH_4OAc, pH 6.4, and injected into a SupercosilTM LC-18-T column (Supelco). The run was performed at a flow rate of 0.7 mL/min using a linear gradient (50 %) of acetonitrile. The electron spray ionization mass spectrometry with a time-of-flight analyzer was performed using a microTOF Focus system (Bruker Daltronics). The capillary voltage was 4500 V, and the end-plate offset was 500 V (negative mode); the dry temperature was 200 °C, the dry gas flow was 9 L/min, and the nebulizer pressure was 2 bar.

DGC activity measurement with a colorimetric assay

To measure initial velocities of YdeH under various conditions, enzymatic activity was measured with the Baykov assay [106]. This coupled spectrophotometric assay quantifies the amount of pyrophosphate, the by-product of the cyclization reaction. Therefore, the reaction mixture contained YdeH in 50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 5 mM $MgCl_2$, 500 milliunits/mL pyrophosphatase from bakers yeast (Sigma) and GTP. In order to stop the reaction at different time points, a 100 μL sample was transferred into 900 μL phosphate detection solution (0.55 M H_2SO_4, 0.46 mM malachite green, 2.66 mM $(NH_4)_6Mo_7O_{24}$, 0.04 % Tween) and the absorption was measured at 630 nm 10 min after incubation. The production of phosphate was quantified with a phosphate calibration curve.
To determine K_M and v_{max}, 2 μM zinc-free YdeH was incubated with different concentrations of GTP (5-100 μM). The measured initial velocities were plotted against GTP concentrations and fitted with a simple Michaelis-Menten curve. Product inhibition of YdeH was measured by mixing 1 μM zinc-free YdeH with different c-di-GMP concentrations (0-2 mM) and 100 or 500 μM GTP. The measured initial velocities were plotted against c-di-GMP concentrations and fitted.
The K_I of the zinc binding was measured with a competition experiment of EDTA and YdeH using the activity measurement with the Baykov assay as readout. Therefore, 0.2 μM

zinc-free YdeH was mixed with 0.2 μM $ZnCl_2$ and different EDTA concentrations (0.1 nM - 1 mM). The activity was plotted against the EDTA concentration and the data were fitted with an exact mathematical model [107] describing competitive binding of EDTA and YdeH to zinc.
The influence of oxidizing conditions on the activity of YdeH was tested with the Baykov assay. For this, 1 μM YdeH and 1 μM $ZnCl_2$ were incubated for 10 or 30 min with 5, 50, 500, and 5000 μM of the following oxidizing agents: iodacetamide, N-ethylmaleimide, H_2O_2, KSCN, $NaAsO_2$, 4-chloromercuribenzoic acid, oxidized glutathione, Paraquat, pyocyanin, pyocianin/NADH, phenylselenyl chloride and oxidized glutathione/glutathione reductase from bakers yeast (1 U, Sigma). Activity measurement was started by adding 100 μM GTP.

2.3.4 Deacetylase activity assay

Preparation of membrane fragments containing PgaC and PgaD was developed by S. Steiner. Therefore, PgaC and PgaD were produced using the strain AB1638 transformed with the plasmid pCD-3xF (S. Steiner). Cells were cultivated at 30 °C in LB-medium supplemented with ampicillin (100 mg/mL) until the OD_{600} had reached 0.2. Gene expression was induced by adding 0.2 % L-arabinose. After 5 h of incubation, cells were harvested by centrifugation (6800 g, 10 min, 4 °C) and the pellets were frozen at −20 °C. Prior to lysis, the frozen cell pellets from 5 L culture were thawed and resuspended in 50 mL 50 mM HEPES, pH 7.5, 50 mM NaCl. After addition of DNAse (2.5 mg/mL), cells were disrupted with a French press at 15 000 psi. Two centrifugation steps (8600 g, 10 min, 4 °C followed by 200 000 g, 60 min, 4 °C) allowed the elimination of unbroken cells and subsequent collection of the total membrane fraction that was resuspended in 20 mL 50 mM HEPES, pH 7.5, 50 mM NaCl. After an additional centrifugation step (200 000 g, 60 min, 4 °C) the membrane fraction was resuspended in 1.25 mL 50 mM HEPES, pH 7.5, 50 mM NaCl and stored at −80 °C.

For the enzymatic reaction, 1 μM PgaB was incubated for 16 h with 4 μL membranes containing PgaC and PgaD and with 2 mM UDP-GlcNAc in 50 mM HEPES, 50 mM NaCl. Subsequently, the proteins were precipitated by heating the sample to 99 °C for 15 min followed by centrifugation (20 800 g, 10 min, 4 °C). The cleared supernatants were analyzed for PgaB deacetylase activity by the detection of acetate with a coupled assay based on enzyme acetate kinase. As readout, the oxidation of NADH to NAD^+ was measured. The measurements were performed using the acetic acid kit (K-ACETAK) from Megazyme. Therefore, 143 μL supernatant were mixed with 138 μL R1 (containing 4 mL bottle 1, 4 tablets from bottle 2, 16 mL water) and 19 μL R2 (1 mL bottle 3, 1.75 mL water) in a flat bottom microtiter plate (falcon) and incubated for 10 min at room temperature before the absorption at 340 nm was measured. The assay was quantified with an acetate calibration curve.

2.3.5 Analysis of complex formation by gel filtration

2.86 nmol $PgaA^{TPR}$ were incubated with 2.86 or 8.56 nmol PgaB in 0.5 mL complex-buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl) for 1 h at room temperature. In other samples, 1 mM $MgCl_2$ or 1 mM GlcNAc was added to the protein mixture. All samples were analyzed for complex formation by gel filtration using a Superdex-200 10/30 column (GE Healthcare) and complex-buffer.

2.3.6 Reductive lysine methylation

1 mg/mL PgaB or YdeH in PgaB-GF buffer (50 mM HEPES, pH 7.5, 50 mM NaCl) was mixed with 20 mM dimethylamine-borane complex and 40 mM formaldehyde and incubated at 4 °C in the dark. After 1 h, 20 mM dimethylamine-borane complex and 40 mM formaldehyde was added again, mixed and incubated again for 1 h at 4 °C in the dark. The reaction was quenched by adding 100 mM Tris-HCl, pH 8.0. The sample was filtered, concentrated to 500 μL and purified from aggregates by gel filtration. For PgaB, a Superdex-200 10/30 column (GE Healthcare) and PgaB-GF buffer was used. YdeH was purified using a HiLoad-16/60-Superdex-75 prep grade column (GE Healthcare) and YdeH-GF buffer (20 mM Tris-HCl, pH 7.6, 150 mM NaCl, 50 mM L-glutamic acid, 50 mM L-arginine).

2.3.7 Limited proteolysis

1 mg/mL PgaB was incubated with 5 ng/mL or 1 ng/mL trypsin, proteinaseK or pepsin in PgaB-GF buffer (50 mM HEPES, pH 7.5, 50 mM NaCl) at room temperature. After different time points of 1, 2, 5, 15, 30, 60, 120 and 240 min, a 10 μL sample was taken and mixed with 1 μL 10 mM protease inhibitor 4-(2-aminoethyl)-benzensulfonylfluorid (AEBSF). The samples were loaded on an SDS-PAGE to follow the proteolytic digestion of PgaB. For large scale digestion, 20 mg PgaB were incubated with 0.1 mg trypsin in 20 mL PgaB-GF buffer for 10 min at room temperature before proteolysis was stopped by adding 1 mM AEBSF. The sample was concentrated to 3 mL and purified by gel filtration using a HiLoad-26/60-Superdex-200 prep grade column (GE Healthcare) and PgaB-GF buffer.

2.4 X-ray crystallography

2.4.1 Protein crystallization

For protein crystallization, the vapor diffusion method was used. Initial screening was performed with commercial screens in 96-well plates using the sitting drop method. For each drop, 0.2 μL protein solution was mixed with an equal amount of reservoir buffer. The reservoir volume was 75 μL. Afterwards the trays were sealed and stored at a constant temperature of 20 °C.

Initial crystallization hits were refined by altering pH, precipitant or salt concentration, protein concentration and temperature. Refinement was carried out in 24-well plates using hanging drops on siliconized cover slips, which were sealed with grease. Protein and reservoir solutions were mixed, usually in a 1:1 ratio and the drop sizes varied between 2 μL and 4 μL. The drops were equilibrated against 500 μL reservoir buffer.
In order to get complex crystals of YdeH with bound product or substrate analog, cocrystallization experiments were performed. Therefore, the protein solution was mixed with the corresponding compound and used in subsequent initial screening experiments.

2.4.2 Data collection

For measurements under cryogenic conditions, crystals were mounted in cryo loops and directly plunged in liquid nitrogen. Prior to freezing, YdeH^{CZB} crystals were successively placed in reservoir buffer supplemented with 10 and 20 % glycerol, using soaking times of 5 sec per buffer. YdeH^{GGDEF} crystals were successively soaked in reservoir solution containing 10 and 15 % glycerol. YdeH crystals were directly frozen from mother liquor without using any additional cryoprotectant.
Measurements using synchrotron radiation were performed at beamline PXIII at the Swiss Light Source (Villigen, Switzerland). Single wavelength anomalous diffraction data was collected on a single selenomethionine containing crystal at the selenium absorption peak, which was determined from a fluorescence scan. For processing and scaling, the XDS package [108] or MOSFLM/SCALA [109, 110] were used.

2.4.3 Structure determination

Structure determination of YdeH^{CZB}

Phase determination of YdeH^{CZB} was done with the single anomalous data method (SAD) using the program autoSHARP [111]. Automated model building was carried out with ARP/wARP [112]. Further model building was performed manually with Coot [113] and refinement was performed with REFMAC [114] using TLS and NCS restraints. Model geometry was analyzed using the programs MolProbity [115] and RAMPAGE [116].

Structure determination of YdeH^{GGDEF}

Phase determination of YdeH^{GGDEF} was done by molecular replacement. The program Phaser was used with chain A of the GGDEF domain of PleD (PDB: 1w25) as search model. After rigid body refinement with REFMAC [114], the model was rebuilt using AutoBuild implemented in Phenix [117]. Refinement was done with REFMAC using one TLS group. Water molecules were placed automatically using Phenix AutoBuild and manually checked

in Coot [113]. Model geometry was analyzed, using the programs MolProbity [115] and RAMPAGE [116].

Structure determination of YdeH

Phase determination was done with molecular replacement using the program Phaser [118] with the CZB domain and the GGDEF domain of YdeH as search models. The model was manually rebuilt in Coot [113] and was refined with BUSTER [119] using NCS restrains and two TLS groups per monomer as well as the two individual domains as target restrains in later stages of refinement. Model geometry was analyzed, using the programs MolProbity [115] and RAMPAGE [116].

2.4.4 Structure analysis and bioinformatics

Secondary structure was assigned using DSSP [120] and STRIDE [121]. Protein interfaces were analyzed with the PISA server [122]. Structure superpositions were performed with SSM [123] and TOPP [124]. Illustrations were made in Pymol (DeLano Scientific). Search for structural homologous was done with the Dali server [125], and a search for binding motifs was performed with PDBeMotif [126]. Structure based sequence alignment was done using the programm ALAdeGAP [127] and illustrated with WebLogo [128] or as conservation mapping with the program ProtSkin [129]. A model of TlpD was created with MODELLER [130] based on the structural sequence alignment with $YdeH^{CZB}$.

3 Zinc dependent regulation of the diguanylate cyclase YdeH from *E.coli*

3.1 Results and Discussion

3.1.1 Design and cloning of YdeH constructs

The plasmid pET28a/YdeH containing the complete sequence of *ydeH* with a C-terminal Histag was kindly provided by Dr. A. Böhm, Biozentrum Basel. Several constructs comprising the individual domains of YdeH were designed for the study of their biochemical properties and to apply them in crystallization attempts. Therefore structures of GGDEF domains in the PDB were analyzed. In all cases, three amino acids in front of the β0, β0'-hairpin were visible in the structures. Thus the construct of the GGDEF domain of YdeH was designed to start at this position corresponding to amino acid Ile127. The construct of the N-terminal CZB domain contained the first 126 amino acids ending with Thr126. These gene fragments were cloned in the pET28a vector as described in section 2.1.

In order to enhance the success of crystallization, surface entropy reduction mutants of YdeH were generated [131, 132]. To identify sites that are most suitable for mutation, the SERp server was used [133]. This server found three sites. The first site contained residues Glu259 and Glu260, the second site included Lys40 and Glu42 and the third site consisted of Lys3 and Lys4. Based on this prediction the following point mutants were designed: E259A-E260A, E259A-E260T, E259T-E260T, E259N-E260N, K40A-K42A and YdeHΔ1-4. The mutations were introduced by site-directed mutagenesis using the QuikChange-method with the pET28a/YdeH plasmid serving as starting vector.

In order to conduct functional studies, mutations of amino acids that may be involved in activation or inhibition of YdeH were designed and generated by site direct mutagenesis, using the Quik-Change-method. All constructs and mutations are listed in Table 3.1.

Table 3.1: YdeH constructs used in this thesis.

Name	Feature
pET28a/YdeH[a]	full-length (1-296)
pET28a/YdeH^{CZB}	N-terminal CZB domain containing the residues 1-126
pET28a/YdeH^{GGDEF}	C-terminal GGDEF domain containing the residues 127-296
pET28a/YdeH-E259A-E260A	surface entropy reduction mutant
pET28a/YdeH-E259A-E260T	surface entropy reduction mutant
pET28a/YdeH-E259T-E260T	surface entropy reduction mutant
pET28a/YdeH-E259N-E260N	surface entropy reduction mutant
pET28a/YdeH-K40A-E42A	surface entropy reduction mutant
pET28a/YdeH-Δ1-4	surface entropy reduction mutant
pET28a/YdeH-C52A	zinc-binding mutant
pET28a/YdeH-R56A	possible I_s-site mutant
pET28a/YdeH-R73A	possible I_s-site mutant
pET28a/YdeH-R87A	possible I_s-site mutant
pET28a/YdeH-R271A-Y274A	possible I_s-site mutant
pET28a/YdeH-D174R	GGDEF dimer interface mutant
pET28a/YdeH-D174R-R140D	GGDEF dimer interface mutant

All constructs contain a C-terminal Histag.
[a] Kindly provided by Dr. A. Böhm, Biozentrum Basel.

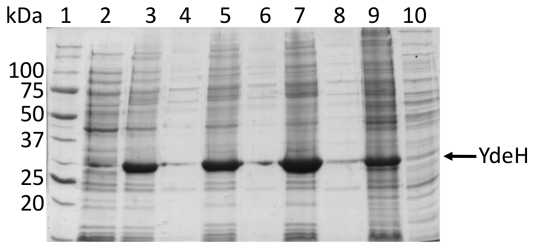
3.1.2 Expression and Purification

Expression of YdeH

Expression tests of YdeH were performed by using different *E. coli* strains at two different temperatures of 37 °C and 18 °C. Expression was induced with 1 mM IPTG. In most of the tested conditions, YdeH was expressed highly in inclusion bodies. The best soluble expression was achieved in the Rosetta(DE3) strain at 37 °C using 4 h of induction. Exemplarily an expression gel is shown in Figure 3.1.
For protein production, bacteria was cultivated on a 2 L or 5 L scale with the conditions established in the expression tests.

Purification of YdeH

YdeH was purified to homogeneity using a two-step purification procedure, consisting of Ni-affinity and size-exclusion chromatography. After cell lysis with French press, Ni-affinity chromatography using a Histrap column was performed. A chromatogram of a typical run is shown in Figure 3.2 and the corresponding SDS-PAGE is shown in Figure 3.3. YdeH is eluting at high imidazole concentrations of about 280 mM, which results in almost pure protein. After Ni-chromatography, YdeH was subjected to gel filtration as a final purification step. Gel filtration was performed in order to remove aggregates, although the protein was rather



lane 1: marker
lane 2: before induction
lane 3: 2 h after induction
lane 4: 2 h after induction, soluble fraction
lane 5: 4 h after induction
lane 6: 4 h after induction, soluble fraction
lane 7: 6 h after induction
lane 8: 6 h after induction, soluble fraction
lane 9: 20 h after induction
lane 10: 20 h after induction, soluble fraction

Figure 3.1: SDS-PAGE of the expression test of YdeH in the pET28a vector in Rosetta(DE3) cells. Expression temperature was 37 °C and induction was carried out with 1 mM IPTG. Same amount of cells were loaded in each line.

pure already. YdeH ran as one single peak at an elution volume of 60.5 mL (Figure 3.4 and Figure 3.5). According to a calibration curve for the present column, this corresponds to a molecular mass of 71 kDa. Assuming the mass of one protein chain of 34 kDa, this clearly indicates that YdeH is a dimeric protein. The final yield was about 5 mg/L culture. To achieve a protein concentration above 0.8 mg/mL, the addition of arginine and glutamic acid to all the buffers was essential to avoid protein aggregation [134].

For YdeH mutants, the same purification protocol was used. However not all mutants were expressed solubly, therefore the expression temperature was reduced to 18 °C and the cells were induced for 20 h. The elution profiles of the gel filtration were identical for all soluble expressed mutants, which indicates that the introduced mutations had no effect on the fold of YdeH. The individual protein yields of YdeH mutants obtained after purification are listed in Table 3.2.

Expression and purification of individual domains of YdeH

The protocol for expressing the individual domains of YdeH was adapted from the expression of the full-length YdeH. Rosetta(DE3) cells grown at 18 °C for 20 h were chosen, which yielded soluble protein in good amounts for both constructs (pET28a/YdeH^{CZB}, pET28a/YdeH^{GGDEF}). For purification, the same procedure as for full-length YdeH was used, consisting of Ni-affinity chromatography followed by gel filtration. The gel filtration chromatograms for both constructs are shown in Figure 3.6 and the corresponding SDS gels are depicted in Figure 3.7. YdeH^{CZB} and YdeH^{GGDEF} ran as a single peak at elution volumes of 71.8 mL and 78.7 mL, respectively. According to a calibration curve for the present column, this corresponds to a molecular mass of 36.8 kDa for the CZB domain. The mass of 14.5 kDa per one protein chain indicates that the CZB domain is dimeric. For the GGDEF domain, a molecular mass of 24.6 kDa was determined with the calibration curve, which

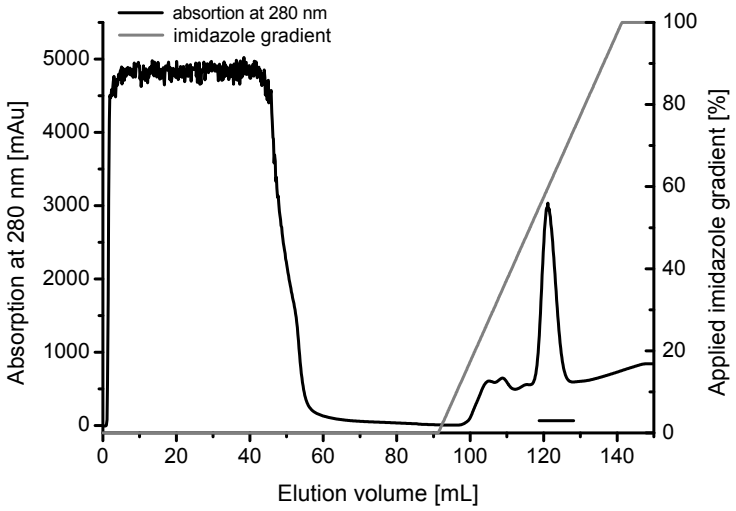
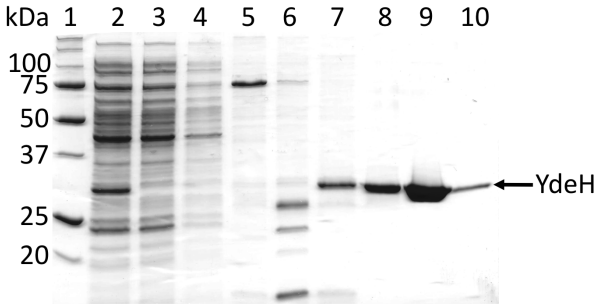


Figure 3.2: Chromatogram of YdeH purification by Ni-affinity chromatography using a 5 mL HisTrap column. YdeH elutes at an imidazole concentration of 280 mM. Fractions pooled and used for subsequent purification are marked by a line.



lane 1:	marker
lane 2:	soluble fraction of the cell extract
lane 3:	flow through (0-55 mL)
lane 4:	wash (55-90 mL)
lane 5:	fraction 104-105 mL
lane 6:	fraction 108-109 mL
lane 7 -10:	fractions 117-128 mL

Figure 3.3: SDS-PAGE of the YdeH purification by Ni-chromatography. 2 μL protein solution were loaded in lane 2-4 and 10 μL in lane 5-10.

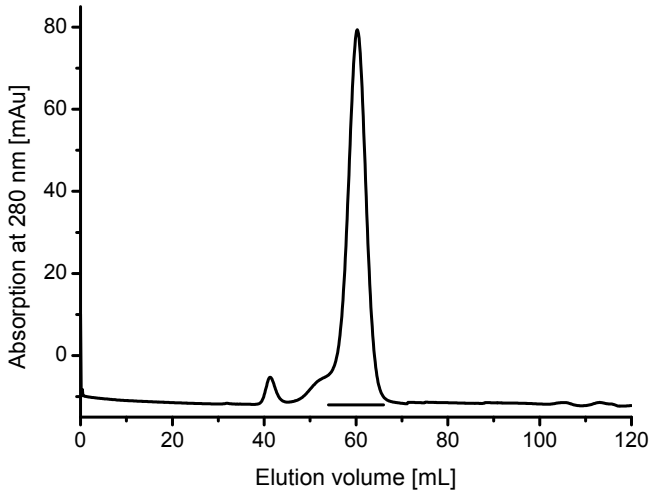
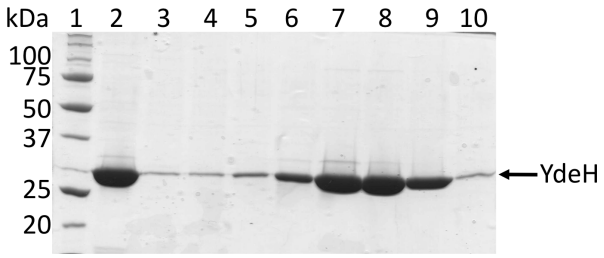


Figure 3.4: Chromatogram of YdeH purification by gel filtration chromatography using a HiLoad-16/60-Superdex-75 prep grade column. YdeH was separated from minor amounts of aggregates, which elute at about 40 mL. Pooled YdeH fractions are marked by a line.



lane 1:	marker
lane 2:	load
lane 3-4:	fractions 40-44 mL
lane 5-10:	wash (55-90 mL)
lane 5-10:	fractions 51-72 mL

Figure 3.5: SDS-PAGE of the YdeH purification by gel filtration. 1 μL protein solution was loaded in lane 2 and 10 μL in lane 5-10.

Table 3.2: Expression and purification results of YdeH mutants.

Mutant	Expression	Yield [mg/L culture]
YdeH wildtype	soluble	5
YdeH-E259A-E260A	not soluble[a]	-
YdeH-E259A-E260T	soluble[a]	0.03
YdeH-E259T-E260T	soluble[a]	0.4
YdeH-E259N-E260N	soluble[a]	0.3
YdeH-K40A-E42A	soluble	2.2
YdeHΔ1-4	soluble[a]	0.2
YdeH-C52A	soluble[a]	7.2
YdeH-R56A	soluble	3.6
YdeH-R73A	soluble	3.0
YdeH-R87A	soluble	1.1
YdeH-R271A-Y274A	soluble	8.4
YdeH-D174R	soluble	6.0
YdeH-D174R-R140D	not soluble[a]	-

[a] Expression performed at 18 °C for 20 h.

is in agreement with the mass of one protein chain of 20.4 kDa. These results show, that the CZB domain is responsible for YdeH dimerization. The final yield were 22 mg/L and 2 mg/L culture for YdeH^{CZB} and YdeH^{GGDEF}, respectively.

Expression and purification of selenomethionine incorporated YdeH^{CZB}

Due to the high expression level and a high number of methionine (4 out of 126 amino acids), the incorporation of selenomethionine was used for phase determination of YdeH^{CZB}. The insertion of methionine was carried out in the methionine-auxotrophic strain B834(DE3) in LeMaster-Medium as described in section 2.2.1. The purification was performed under reducing conditions by adding 0.5 mM TCEP in all buffers. The purification procedure was the same as for wild type YdeH^{CZB} (see section 3.1.2). The obtained protein was pure and the final yield was about 2.5 mg/L culture. To verify the incorporation of selenomethionine, mass spectral analysis was performed. In comparison to wildtype YdeH^{CZB} the difference in mass was 184 Da, which corresponds exactly to four incorporated selenomethionines.

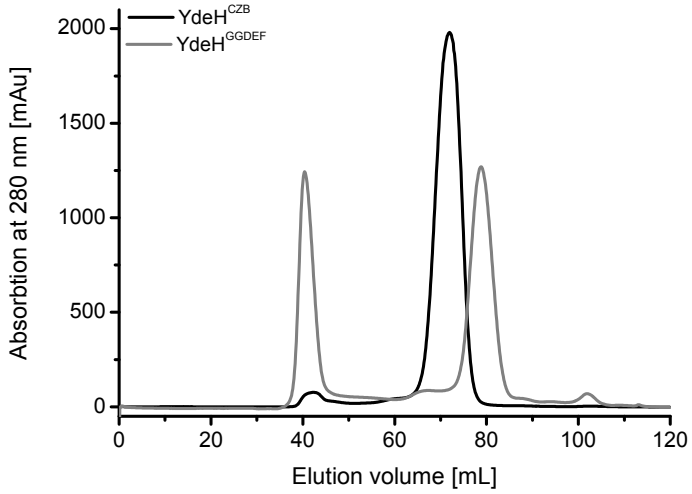
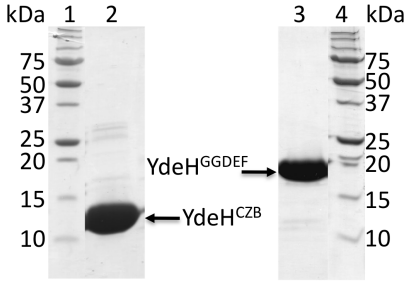


Figure 3.6: Chromatograms of $YdeH^{CZB}$ and $YdeH^{GGDEF}$ purification by gel filtration chromatography using a HiLoad-16/60-Superdex-75 prep grade column. $YdeH^{CZB}$ (black line) and $YdeH^{GGDEF}$ (grey line) were separated from aggregates, which elute at about 40 mL.



lane 1: marker
lane 2: $YdeH^{CZB}$ after gel filtration
lane 3: $YdeH^{GGDEF}$ after gel filtration
lane 4: marker

Figure 3.7: SDS-PAGE of $YdeH^{CZB}$ and $YdeH^{GGDEF}$ after gel filtration. 10 μL protein solution were loaded in each lane.

3.1.3 Crystallization

Crystallization of full-length YdeH

Many attempts were made to obtain crystals of full-length YdeH. For this, commercial crystallization screens were employed at different protein concentrations in 96-well sitting drop set-ups. All trials are listed in Table 3.3. As the first attempts to obtain crystals of wildtype YdeH were not successful, YdeH was transferred to YdeH-GF-buffer without the addition of arginine and glutamate. However, no crystals were growing under such conditions either. Since YdeH has five cysteines, which could crosslink YdeH and form oligomeric species, YdeH was crystallized under reductive conditions by adding 0.5 mM TCEP to all buffers during protein purification. However, no difference was observed in gel filtration experiments and in precipitation behavior in the crystallization setups.

Thereupon, cocrystallization was tested with the non-reactive substrate analogs ddGTP, GTPαS and GDP in the presence of magnesium and with GTP and calcium, which inhibits YdeH. In addition, crystallization set-ups in the presence of the reaction product c-di-GMP and magnesium were performed, since most of the crystallized DGCs so far have c-di-GMP bound to the I-site. In these structures, c-di-GMP is involved in domain crosslinking and thus, stabilizes the proteins by mediating crystal contacts. However, in all cocrystallization experiments, no crystals could be observed.
To mimic the transition state of the catalytic reaction, an inhibitor (TSA) was designed. In this state, the formation of the first bond between the two GTP molecules is already generated, whereas the second bond is not yet formed. The inhibitor TSA should crosslink the two active sites of the two GGDEF domains and fix them in one conformation. This arrangement should provide more insights into the mechanism of the catalysis. This inhibitor was synthesized by YdeH from ddGTP and GTPαS as it is described in appendix **??**. The cocrystallization experiments with this inhibitor did not yield any crystals.

Surface entropy reduction mutants were tried to allow crystallization. In these mutants large hydrophilic residues were changed to smaller residues like alanine, which leads to a local reduction of conformational entropy and generates contact-forming, conformationally homogenous surface patches [131, 132]. As the surface charge was reduced along with the solubility, not all mutants gave enough protein for crystallization experiments. The mutants K40A-E42A, YdeHΔ1-4, E259T-E260T and E259N-E260N could be tested in cocrystallization experiments with substrate analogs and product as it is listed in Table 3.3. However, no protein crystals could be obtained from these mutants in the different crystallization set-ups.

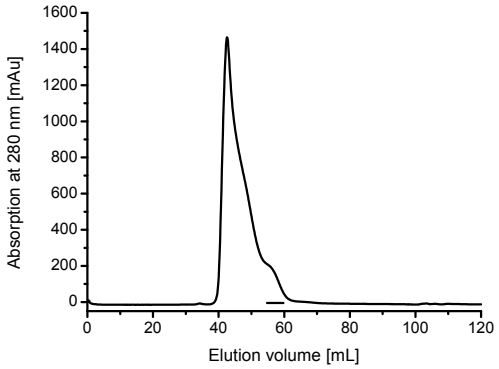


Figure 3.8: Chromatogram of YdeH after methylation of surface lysines. A HiLoad-16/60-Superdex-75 prep grade column was used. Most of the YdeH elutes as aggregates at about 45 mL. The pooled fractions, which were used for crystallization experiments, are marked by a black line.

A method, which is based on a similar principle, is the methylation of surface lysines [135]: dimethylated lysines will reduce their interaction with solvent molecules, thereby forcing them to adopt a more ordered conformation that may facilitate crystallization. Fore this, YdeH was methylated as described in section 2.3.6 and purified from aggregates by gel filtration. Because of the rather harsh conditions for the methylation, the loss of protein is high, even by performing the experiment at 4 °C. Not surprisingly, the major part of YdeH elutes as aggregates in gel filtration experiments (Figure 3.8). Although the peak containing YdeH was not well separated from the aggregates, the fractions marked with a line in the chromatogram were further used for crystallization experiments. However, no crystals of methylated YdeH could be obtained.

After the structure of the CZB domain of YdeH was solved and showed the presence of a zinc ion (see section 3.1.6), YdeH was tried to crystallize in the presence of zinc in order to obtain a homogenous sample, which should crystallize more easily. This did not result in protein crystals as well.
In addition, a zinc binding mutant C52A was tried to crystallize. For this, initial screening was performed in the presence of zinc as well as cocrystallization experiments with GTPαS and c-di-GMP as it is listed in Table 3.3. In the presence of GTPαS, crystals were obtained in two similar conditions containing 200 mM magnesium acetate and 20 % PEG 3350 or 200 mM calcium acetate and 20 % PEG 3350 (Figure 3.9). They grew after three days and reached their final size after about three weeks. These conditions were refined, but the crystals could not be reproduced. However, the crystals from the initial screen were large enough to use them in diffraction experiments at the synchrotron.

Table 3.3: Crystallization trails for full-length YdeH.

YdeH	Concentration [mg/mL]	Additives	Used screens
wildtype	2.5	-	Crystal Screen, Classics light Suite, PEGs Suite, MPD Suite, $AmSO_4$ Suite, Wizard I+II
wildtype	4	-	Anions Suite, Cations Suite
wildtype	5.4	-	Crystal Screen, Classics light Suite, PEGs Suite, Wizard I+II
wildtype	10	-	Crystal Screen, Classics light Suite, PEGs Suite, MPD Suite, $AmSO_4$ Suite, Wizard I+II, Anions Suite, Cations Suite
wildtype	10	without Arg and Glu in the gel filtration buffer	Crystal Screen, PEGs Suite, Index, SaltRx
wildtype	4	100 mM LiCl	PEGs Suite, $AmSO_4$ Suite
wildtype	8	200 μM $ZnCl_2$	PACT premier, JCSG-*plus*, Structure screen, MbClass Suite, Index
wildtype	4	3 mM ddGTP, 5 mM $MgCl_2$	Crystal Screen, Classics light Suite, PEGs Suite, MPD Suite, $AmSO_4$ Suite, Wizard I+II, Anions Suite, Cations Suite
wildtype	4	3 mM GTPαS, 5 mM $MgCl_2$	Crystal Screen, Classics light Suite, PEGs Suite, MPD Suite, $AmSO_4$ Suite, Wizard I+II, Anions Suite, Cations Suite
wildtype	4	3 mM GTP, 5 mM $CaCl_2$	Crystal Screen, Classics light Suite, PEGs Suite, MPD Suite, $AmSO_4$ Suite, Wizard I+II, Anions Suite, Cations Suite
wildtype	4	5 mM $MgCl_2$	Crystal Screen, Classics light Suite, PEGs Suite, MPD Suite, $AmSO_4$ Suite, Wizard I+II, Anions Suite, Cations Suite
wildtype	5	5 mM c-di-GMP, 10 mM $MgCl_2$	Crystal Screen, Classics light Suite, PEGs Suite, MPD Suite, $AmSO_4$ Suite, Wizard I+II, Anions Suite, Cations Suite

continued on next page

continued from previous page

YdeH	Concentration [mg/mL]	Additives	Used screens
wildtype	10	5 mM c-di-GMP, 10 mM $MgCl_2$	Crystal Screen, Classics light Suite, PEGs Suite, MPD Suite, $AmSO_4$ Suite, Wizard I+II, Anions Suite, Cations Suite
wildtype	10	3 mM ddGTP, 5 mM $MgCl_2$	Crystal Screen, PEGs Suite, Index, SaltRx
wildtype	10	3 mM GTPαS, 5 mM $MgCl_2$	Crystal Screen, PEGs Suite, Index, SaltRx
wildtype	10	3 mM GTP, 5 mM $CaCl_2$	Crystal Screen, PEGs Suite, Index, SaltRx
wildtype	10	3 mM GDP, 5 mM $MgCl_2$	Crystal Screen, PEGs Suite, Index, SaltRx
wildtype	8	5 mM Inhibitor TSA, 10 mM $MgCl_2$	Crystal Screen, PEGs Suite, Index, SaltRx
wildtype	8	0.5 mM TCEP	Crystal Screen, Classics light Suite, MPD Suite, Wizard I+II, SaltRx
wildtype	8	0.5 mM TCEP, 3 mM GTPαS, 5 mM $MgCl_2$	Crystal Screen, Classics light Suite, MPD Suite, Wizard I+II, SaltRx
wildtype	8	0.5 mM TCEP, 5 mM c-di-GMP, 5 mM $MgCl_2$	Crystal Screen, Classics light Suite, MPD Suite, Wizard I+II, SaltRx
methylated	5	-	Crystal Screen, PEGs Suite, Index, SaltRx
methylated	8	-	Crystal Screen, PEGs Suite, Index, SaltRx
K40A-E42A	5	-	Crystal Screen, Classics light Suite, PEGs Suite, MPD Suite, $AmSO_4$ Suite, Wizard I+II, Anions Suite, Cations Suite, Index, SaltRx, MbClass Suite
K40A-E42A	8	-	Crystal Screen, Classics light Suite, PEGs Suite, MPD Suite, $AmSO_4$ Suite, Wizard I+II, Anions Suite, Cations Suite, Index, SaltRx, MbClass Suite

continued on next page

continued from previous page

YdeH	Concentration [mg/mL]	Additives	Used screens
K40A-E42A	5	5 mM c-di-GMP, 5 mM $MgCl_2$	Crystal Screen, PEGs Suite, $AmSO_4$ Suite, Index, SaltRx
K40A-E42A	8	5 mM c-di-GMP, 5 mM $MgCl_2$	Crystal Screen, PEGs Suite, $AmSO_4$ Suite, Index, SaltRx
YdeHΔ1-4	5	-	Crystal Screen, Classics light Suite, PEGs Suite, MPD Suite, $AmSO_4$ Suite, Wizard I+II, Anions Suite, Cations Suite, Index, SaltRx, MbClass Suite
E259T-E260T	6.4	-	Crystal Screen, PEGs Suite, Wizard I+II, SaltRx
E259N-E260N	5	-	Crystal Screen, Wizard I+II, Index
E259N-E260N	8	-	Crystal Screen, Wizard I+II, Index
E259N-E260N	5	3 mM GTPαS, 5 mM $MgCl_2$	Crystal Screen, Wizard I+II, Index
C52A	8	200 μM $ZnCl_2$	PACT premier, Structure screen, Crystal Screen, PEGs Suite
C52A	8	200 μM $ZnCl_2$, 3 mM GTPαS, 5 mM $MgCl_2$	PACT premier, Structure screen, Crystal Screen, PEGs Suite
C52A	8	200 μM $ZnCl_2$, 5 mM c-di-GMP, 5 mM $MgCl_2$	PACT premier, Structure screen, Crystal Screen, PEGs Suite

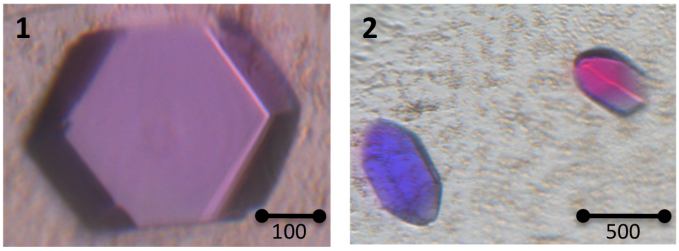


Figure 3.9: Crystals of YdeH-C52A obtained in the initial screening. On the left, crystals were grown in 200 mM magnesium acetate and 20 % PEG 3350 and on the right in 200 mM calcium acetate and 20 % PEG 3350. In both conditions, 3 mM GTPαS, 5 mM $MgCl_2$ and 200 μM $ZnCl_2$ was present.

The difficulty of getting full-length YdeH crystals might be due to variability in the domain arrangement. This flexibility may be important for catalysis, but prevents the protein to be present in one conformation, which usually facilitates crystallization. Recombinant YdeH is active in solution (section 3.1.10) and elutes as a single peak in gel filtration. The same results could be obtained after storing YdeH at 4 °C for 20 days. This indicates, that full-length YdeH is indeed properly folded and stable. In order to avoid the variability between the two domains of YdeH, isolated domains, $YdeH^{CZB}$ and $YdeH^{GGDEF}$, were used to perform crystallization trials.

Crystallization of $YdeH^{CZB}$

Initial crystallization trials with $YdeH^{CZB}$ were set up in 96-well sitting drop trays. The commercial screens Crystal Screen, Classic light Suite, SaltRx and PEGs Suite were tested at protein concentrations of 10, 20 and 36 mg/mL. Crystals were obtained under a variety of conditions ranging from pH 4.6 to 8.5. Most of the conditions contained PEG from size of 2000 to 8000 and an ammonium salt. The obtained crystals are shown in Figure 3.10 and the corresponding conditions are listed in Table 3.4.
In order to get bigger crystals, all the initial conditions were refined by altering pH, PEG or salt concentration, protein concentration and temperature. The diffraction quality of the refined crystals was tested on the home source. The refined conditions of hit 1, 4, 6, and 12 gave the most promising results and were used for diffraction experiments at the synchrotron. The best results were obtained in the condition containing 200 mM ammonium tartrate, 12.5 % PEG 3350 with a protein concentration of 34 mg/mL at 20 °C.

Crystallization experiments with the selenomethionine incorporated protein were set up in 24-well plates using the best four conditions (1, 4, 6 and 12) obtained for the wildtype $YdeH^{CZB}$. In all four conditions, crystals of SeMet $YdeH^{CZB}$ could be obtained. The best

Table 3.4: Initial crystallization conditions of YdeH^{CZB}.

Hit	Protein concentration [mg/mL]	Condition
1	20	200 mM ammonium acetate, 100 mM sodium citrate, pH 5.6, 30 % PEG 4000
2	36	200 mM ammonium acetate, 100 mM sodium citrate, pH 5.6, 15 % PEG 4000
3	10, 20	200 mM ammonium acetate, 100 mM sodium acetate, pH 4.6, 15 % PEG 4000
4	20, 36	200 mM sodium acetate, 100 mM sodium cacodylate, pH 6.5, 15 % PEG 8000
5	36	800 mM sodium citrate
6	10	200 mM ammonium tartrate, 20 % PEG 3350
7	10, 20	200 mM ammonium fluoride, 20 % PEG 3350
8	10, 20	200 mM ammonium acetate, 20 % PEG 3350
9	20	200 mM ammonium nitrate, 20 % PEG 3350
10	20, 36	200 mM ammonium formiate, 20 % PEG 3350
11	36	200 mM ammonium citrate, 20 % PEG 3350
12	20	100 mM MES, pH 6.5, 25 % PEG 2000 MME
13	36	100 mM HEPES, pH 7.5, 700 mM sodium citrate
14	20	100 mM Bis-Tris propane, pH 7.0, 3.5 M sodium formiate
15	10, 20	100 mM Tris-HCl, pH 8.5, 3.5 M ammonium chloride

diffraction could be achieved from crystals using a protein concentration of 30 mg/mL grown in 200 mM sodium tartrate, 12.5 % PEG 3350 at 20 °C.

Crystallization of YdeH^{GGDEF}

Initial crystallization trials using YdeH^{GGDEF} at a concentration of 5 and 10 mg/mL using the commercial screens Crystal Screen, Index, PEGs Suite and $AmSO_4$ Suite failed to produce any crystal. However, the addition of 5 mM c-di-GMP and 5 mM $MgCl_2$ at a protein concentration of 5 mg/mL yielded crystals in 92 conditions. The crystals grew under a wide pH range from 4.6 to 9.0 using different buffers. Different salts were used and all PEG sizes were represented along with other precipitants. A variety of crystal forms could be observed, such as plates, needles, rods, with most of the hits being octahedral crystals (71). Two conditions, which are listed in Table 3.5 and shown in Figure 3.11, were refined by varying pH and PEG concentration as well as protein concentration. The crystals grown in the refined condition 100 mM Tris-HCl, pH 8.5, 22.5 % PEG 4000 with 5 mg/mL YdeH^{GGDEF} gave the best diffraction in-house and were used for data collection at the synchrotron.

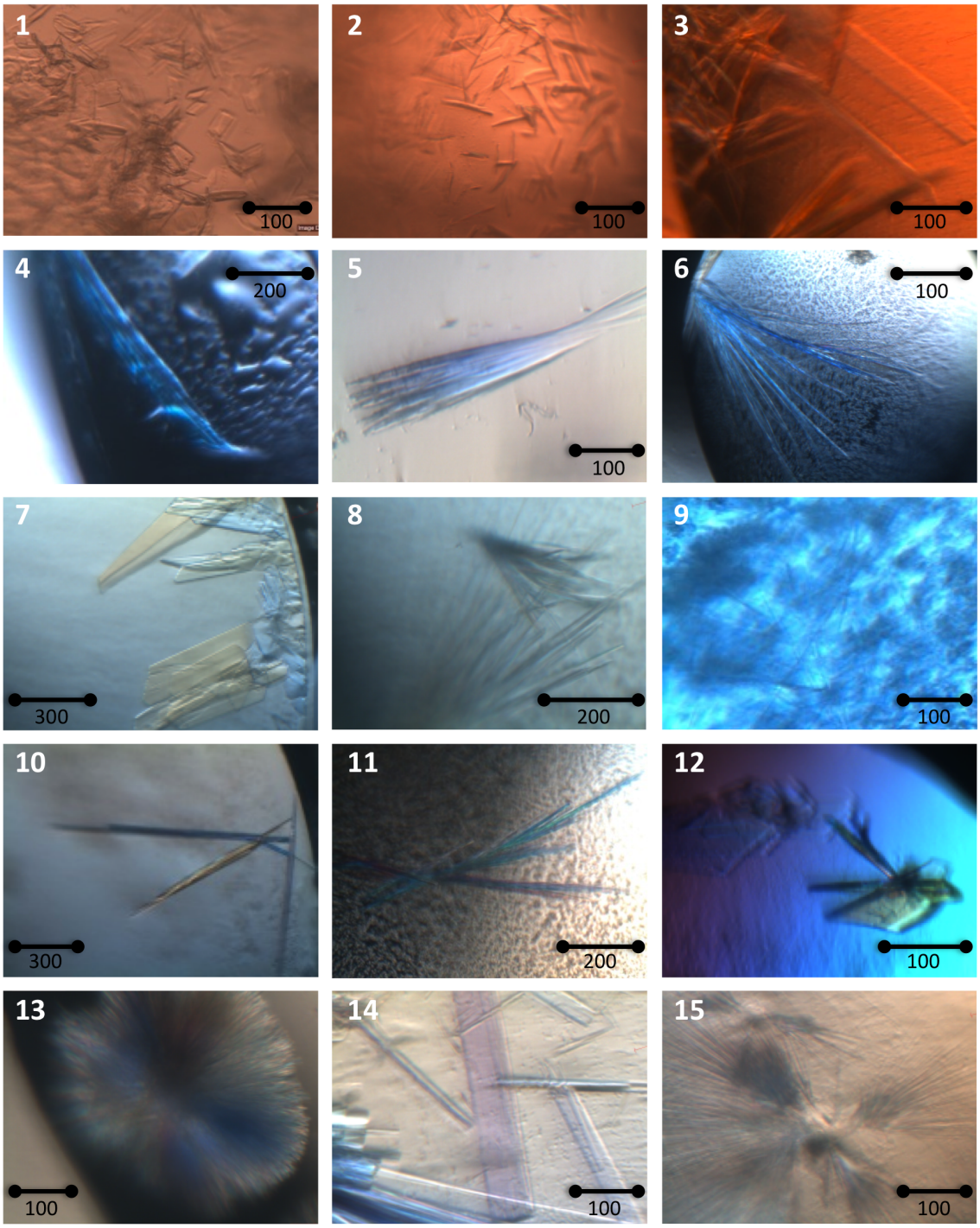


Figure 3.10: Crystals of YdeH^{CZB} obtained in the initial screening. The numbering is the same as in Table 3.3. Dimension are in μm.

Table 3.5: Initial crystallization conditions of YdeH^{GGDEF}.

Hit	Protein concentration [mg/mL]	Additive	Condition
1	5	5 mM c-di-GMP, 5 mM $MgCl_2$	100 mM HEPES, pH 7.5, 30 % PEG 400
2	5	5 mM c-di-GMP, 5 mM $MgCl_2$	100 mM Tris-HCl, pH 8.5, 25 % PEG 4000

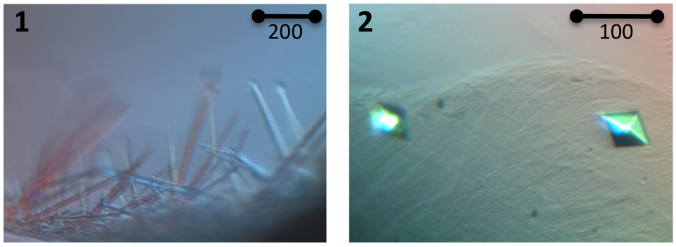


Figure 3.11: Crystals of YdeH^{GGDEF}, which were obtained in the initial screening and were chosen for further refinement. The numbering is the same as in Table 3.5. Dimension are in μm.

3.1.4 Structure determination

Structure determination of YdeH^{CZB}

Data collection

For data collection, a cryoprotection protocol was developed that involved successive transfer of YdeH^{CZB} crystals into reservoir solution supplemented with 10 and 20 % glycerol and were subsequently flash frozen in liquid nitrogen. The soaking time in each drop was kept as short as possible to minimize osmotic stress.
Because of the absence of a search model for molecular replacement, the single anomalous diffraction method (SAD) on selenomethionine incorporated crystals was chosen. The data collection was performed at beamline PXIII at the Swiss Light Source (Villigen, Switzerland). One high-resolution native dataset and a SAD dataset at the peak wavelength of the selenium edge (0.9792 Å) of a SeMet crystal were collected. An X-ray fluorescence scan was performed to determine the correct wavelength for the SAD experiment.
All diffraction data were processed with XDS and scaled with XSCALE [108]. Data collection parameters and statistics are given in Table 3.6. Indexing was straightforward and the crystals belonged to the monoclinic space group C2 with two molecules per asymmetric unit and a solvent content of 48 %. Despite the rather low redundancy of the anomalous data, the signal to noise ratio of the anomalous differences and the correlation of anomalous differences between two randomly chosen subsets were significant up to a resolution of 2.55 Å (Table 3.7).

Table 3.6: Data collection parameters and statistics for the native and SeMet dataset of YdeH^{CZB}.

	Native	SeMet
Wavelength [Å]	0.9792	0.9792
Detector	MarCCD 225	MarCCD 225
Detector distance [mm]	180	160
Space group	C2	C2
Cell axes [Å]	$a = 121.46, b = 60.26, c = 38.46$	$a = 119.07, b = 59.04, c = 37.63$
Angles [°]	$\alpha = \gamma = 90$, $\beta = 99.04$	$\alpha = \gamma = 90, \beta = 99.12$
Resolution [Å]	50.0-2.20 (2.25-2.20)	50.0-2.55 (2.62-2.55)
Observed reflections [a]	57817 (7079)	67492 (4933)
Unique reflections [a]	14022 (1739)	16332 (1159)
Multiplicity	4.1	4.1
R_{merge} [%] [a]	6.9 (39.5)	5.6 (14.2)
$I/\sigma(I)$[a]	14.6 (4.1)	19.1 (10.7)
Completeness [%] [a]	99.7 (100)	99.0 (100)
Wilson B [Å^2]	38.1	35.0
Protomers per ASU	2	2
V_M [$Å^3 \cdot Da^{-1}$]	2.39	2.24
Solvent content [%]	48.5	45.1

[a] Values in parentheses refer to the highest resolution bin.

Table 3.7: Signal-to-noise (S/N) ratio and correlation coefficients (CC) of the anomalous differences in the SAD dataset of YdeH^{CZB}.

Resolution shell [Å]	30.0 - 10.0	10.0 - 6.0	6.0 - 5.0	5.0 - 4.0	4.0 - 3.5	3.5 - 3.1	3.1 - 2.9	2.9 - 2.7	2.7 - 2.6	2.6 - 2.5	2.5 - 2.4
S/N ratio[a]	3.0	3.4	3.0	2.3	1.9	1.7	1.6	1.5	1.3	1.2	1.1
CC [%][b]	89	90	87	79	65	64	58	53	46	38	32

[a]Calculated as the ratio of anomalous differences to their standard deviations.
[b]Correlation coefficient of the anomalous differences between to subsets.

Table 3.8: Heavy atom parameters and SAD phasing statistics for YdeH^{CZB}.

Phasing statistics					
Resolution range [Å]	35.0 - 2.2				
Phasing power					
isomorphous	0.90				
anomalous	2.31				
R_{cullis}					
isomorphous	0.82				
anomalous	0.54				
FOM	0.33				
Heavy atom site	x	y	z	occupancy	B-factor [Å^2]
Se1	82.070	3.676	17.061	0.79	30.30
Se2	77.161	10.584	2.078	0.69	32.49
Se3	92.613	0.644	21.618	0.81	38.53
Se4	85.236	8.599	1.099	0.71	32.66
Se5	69.725	2.042	12.548	0.83	46.49
Se6	96.779	10.625	0.475	0.79	41.31
Se7	72.096	11.095	0.918	0.75	43.41
Se8	74.445	2.004	13.256	0.74	49.72
Zn9	29.236	14.827	36.824	0.87	37.08
Zn10	21.707	27.571	17.020	0.79	41.16

Selenium substructure determination and phase determination by SAD

For heavy atom detection and phasing the program autoSHARP [111] was used. A clear solution with eight selenium atoms was obtained corresponding to the four methionine residues per protomer. All heavy atoms were refined to occupancies close to unity and similar isotropic B factors (Table 3.8), indicating nearly complete incorporation during protein synthesis and a similar degree of order for all SeMet side chains. During refinement of the heavy atom sites two new sites were detected, which were later modeled as zinc ions. The first hint that the additional atoms were zinc was given by the fluorescence scan, in which an anomalous signal was measured at the zinc edge (0.7478 Å). Later this finding was confirmed by the typical coordination and bond lengths of zinc. Finally the program autoSharp performed density modification, which resulted in an electron density map of high quality that was easily interpretable (Figure 3.12).

Model building and refinement

Automatic model building was performed with ARP/wARP [112], that places 195 out of 264 residues in the electron density. Loops were manually built with Coot [113] in the electron density belonging to the protomer B, which was judged to be better defined as protomer A.

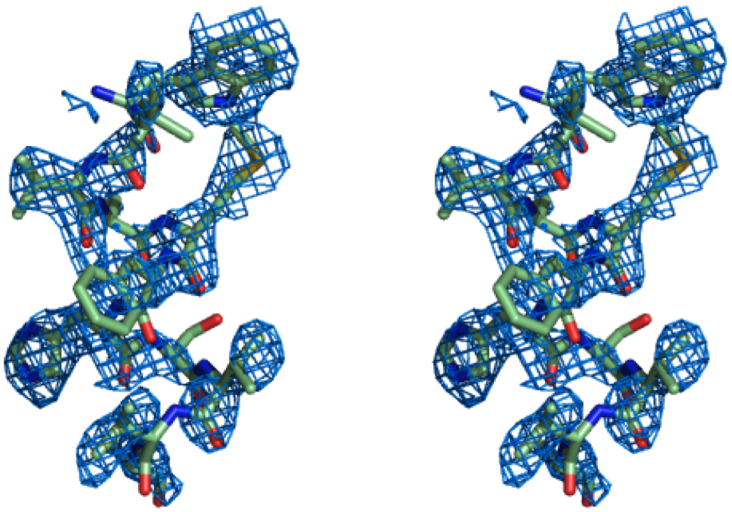


Figure 3.12: Experimental electron density for YdeH^{CZB} after SAD phasing and density modification. Electron density map at 1.5 σ contour level is shown together with residues 25-35 of the final model.

Monomer A was deleted and newly generated from protomer B by application of the NCS operator. This model of the YdeH^{CZB} dimer was subjected to a first round of refinement in REFMAC [114]. Strong NCS restraints were used in early stages of refinement, but loosened later on. Both monomers exhibited strong difference density, which was filled with a zinc ion. In further rounds of rebuilding and refinement, TLS parameterization was added in form of one group per protomer and water molecules were added in Coot. This resulted in a final model of YdeH^{CZB} with R and R_{free} values of 21.1 % and 24.6 %, respectively (Table 3.9).

Validation of the YdeH^{CZB} model

The refined model of YdeH^{CZB} consists of a two-fold symmetric dimer. The electron density was well defined for the whole main chain except for a 14 residue segment (38-51) of chain A as well as for the first four residues of chain A and the first five residues of chain B. Refinement progress was followed by monitoring the values of R and R_{free} and checking for model geometry. The final statistics are given in Table 3.9. The values for the R-factors were reasonable for the resolution of 2.2 Å and the RMS deviations of bond lengths and angles were in the standard range. All main chain torsion angles lay in the allowed regions of the Ramachandran plot, with most of them in the favored region (Figure 3.13).

Table 3.9: Refinement statistics for the YdeH^{CZB} model.

Resolution[Å]	2.2
Space group	C2
R [%]	21.1
R_{free} [%]	24.6
Number of atoms	
protein	1844
water molecules	61
zinc	2
B factors	
main chain A / main chain B	32.3 / 32.7
side chain A / side chain B	33.9 / 34.6
zinc	23.3
water molecules	34.9
RMS deviations	
bond lengths [Å]	0.012
bond angles [°]	1.206
Ramachandran statistics [%]	
in favored regions	98.6
in allowed regions	1.4

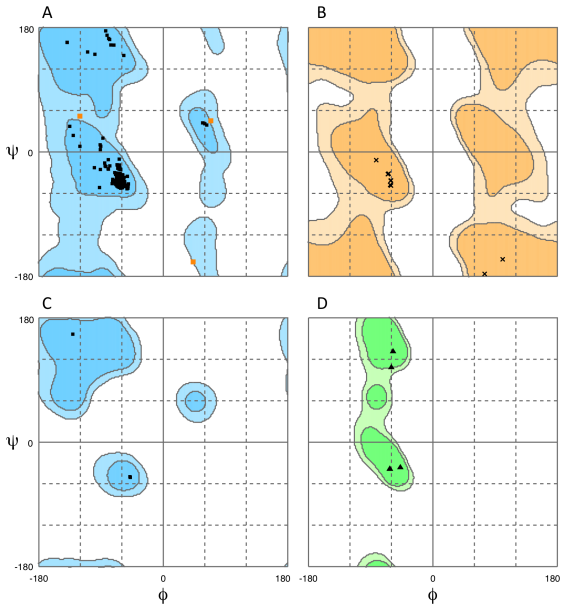


Figure 3.13: Ramachandran diagram for YdeH^{CZB}. Main chain torsion angles were analyzed with RAMPAGE [116] and are indicated for each residue by crosses (glycines) or squares (all others). Dark and light background shading indicates energetically favorable and allowed regions, respectively. (A) General plot for all non-glycine residues. (B) Special plot for glycine residues. (C, D) Plots for pre-proline and proline residues, respectively.

Table 3.10: Data collection parameters and statistics of YdeH^{GGDEF}.

Wavelength [Å]	0.979
Detector	MarCCD 225
Detector distance [mm]	180
Space group	$P6_2$
Cell axes [Å]	$a = b = 79.62, c = 51.21$
Angles [°]	$\alpha = \beta = 90, \gamma = 120$
Resolution [Å]	50.00-1.80 (1.90-1.80)
Observed reflections [a]	216535 (30101)
Unique reflections [a]	17292 (2509)
Multiplicity	12.5
R_{merge} [%] [a]	7.5 (34.6)
$I/\sigma(I)^a$	24.1 (7.6)
Completeness [%] [a]	100 (100)
Wilson B [Å^2]	14.9
Number of protomers per ASU	1
V_M [Å$^3 \cdot Da^{-1}$]	2.30
Solvent content [%]	46.5

[a] Values in parentheses refer to the highest resolution bin.

Structure determination of YdeH^{GGDEF}

Data collection

Crystals of YdeH^{GGDEF} were soaked in reservoir solution containing glycerol as cryoprotectant. The concentration of the glycerol has been increased in two steps from 10 to 15 % before the crystals were flash frozen in liquid nitrogen.
A high-resolution dataset (1.8 Å was collected at the SLS beamline PXIII (Villigen, Switzerland). Data were processed with MOSFLM [109] and scaled with SCALA [110]. The parameters and statistics of the data collection are shown in Table 3.10. The crystal belonged to the space group $P6_2$ with one molecule in the asymmetric unit and a solvent content of 46 %. The analysis of the relevant map selections of the self-rotation function showed no signs for closed rotational NCS.

Phase determination by molecular replacement, model building and refinement

As several structures of GGDEF domains were already available, molecular replacement could be used for phase determination. Different programs (MOLREP[136], Phaser [118]) combined with several different search models (PDB: 1w25, 3i5b, 3ign) were tested. All gave very poor solutions with Z-scores below eight. The best result was achieved with the program Phaser and with chain A of the GGDEF domain of PleD (PDB: 1w25) as a search model. This initial solution had no clashes, a Z-score of 7.8 and a log likelihood of 71.

After a rigid body refinement with REFMAC [114], the model was rebuilt using AutoBuild implemented in Phenix [117]. The programm placed 157 out of 170 residues and improved the model drastically resulting in much better R and R_{free} values of 26 % and 30 % respectively. Refinement was done with REFMAC using one TLS group. In the active and the I-site well defined difference electron density was visible, which could be clearly assigned to c-di-GMP. The c-di-GMP molecule in the A-site sits on a two-fold crystallographic axis, so that only half of the c-di-GMP molecule had to be built in the electron density. A c-di-GMP dimer was placed in the difference electron density close to the I-site. Water molecules were placed automatically using Phenix AutoBuild and manually checked in Coot [113]. Towards the end of refinement, a magnesium ion could be placed in the difference electron density close to c-di-GMP in the I-site. After several rounds of rebuilding in Coot and refinement with REFMAC, the final model resulted in R and R_{free} values of 16.6 % and 20.1 %, respectively (Table 3.11).

Validation of the YdeH^{GGDEF} model

Electron density was well defined for the whole main chain of YdeH^{GGDEF} as well as for a c-di-GMP molecule at the active site, a c-di-GMP-dimer at the I-site and a magnesium ion. Refinement statistics of the YdeH^{GGDEF} model are given in Table 3.11. Refinement progress was followed by monitoring the values of R and R_{free} throughout model building. Stereochemical parameters such as RMS deviations of bond length and angles were in the normal range. The geometry of the YdeH^{GGDEF} model was analyzed using the programs MolProbity [115] and RAMPAGE [116]. All main chain torsion angles were in the most favored or allowed region of the Ramachandran plot (Figure 3.14). The side chain conformations showed no outliers, either.

Structure determination of full-length YdeH

Data collection

As before tested in-house, full-length YdeH crystals didn't need any additional cryoprotectant and were directly flash frozen in liquid nitrogen. A dataset up to 3.9 Å was collected at the SLS beamline PXIII (Villigen, Switzerland). Data were processed with MOSFLM [109] and scaled with SCALA [110]. The parameters and statistics of the data collection are shown in Table 3.12. The crystal belonged to the space group $P6_1$, with two molecules in the asymmetric unit and a solvent content of 57 %.

Table 3.11: Refinement statistics for the YdeH^{GGDEF} model.

Resolution[Å]	1.8
Space group	$P6_2$
R [%]	16.6
R_{free} [%]	20.1
Number of atoms	
protein	1519
water molecules	140
c-di-GMP	115
magnesium	1
B factors	
main chain	16.7
side chain	20.8
c-di-GMP	16.0
magnesium	29.6
water molecules	26.0
RMS deviations	
bond lengths [Å]	0.009
bond angles [°]	1.281
Ramachandran statistics [%]	
in favored regions	99.4
in allowed regions	0.6

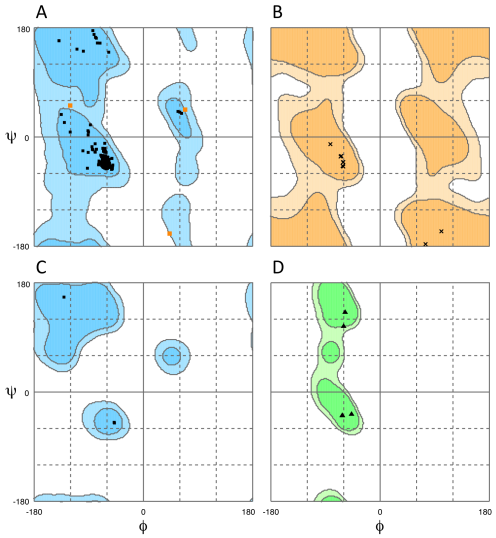


Figure 3.14: Ramachandran diagram for YdeH^{GGDEF}. Main chain torsion angles were analyzed with RAMPAGE [116] and are indicated for each residue by crosses (glycines) or squares (all others). Dark and light background shading indicates energetically favorable and allowed regions, respectively. (A) General plot for all non-glycine residues. (B) Special plot for glycine residues. (C, D) Plots for pre-proline and proline residues, respectively.

Table 3.12: Data collection parameters and statistics for full-length YdeH.

Wavelength [Å]	0.979
Detector	MarCCD 225
Detector distance [mm]	320
Space group	$P6_1$
Cell axes [Å]	$a = b = 102.59, c = 129.16$
Angles [°]	$\alpha = \beta = 90, \gamma = 120$
Resolution [Å]	50.0-3.90 (4.1-3.9)
Observed reflections [a]	92268 (8301)
Unique reflections [a]	7085 (1047)
Multiplicity	13.0
R_{merge} [%] [a]	13.6 (77.4)
$I/\sigma(I)^a$	13.6 (2.7)
Completeness [%] [a]	99.9 (99.9)
Number of protomers per ASU	2
V_M [$Å^3 \cdot Da^{-1}$]	2.89
Solvent content [%]	57.4

[a] Values in parentheses refer to the highest resolution bin.

Phase determination by molecular replacement, model building and refinement

As the structures of the two individual domains of YdeH were already solved, molecular replacement was used for phase determination. This was done with the program Phaser [118] with the CZB domain and the GGDEF domain of YdeH as search models. Two molecules of the CZB domain and two molecules of the GGDEF domain were placed by the procedure in the asymmetric unit in a sensible arrangement. By placing more and more of the domains, the Z-score and the log likelihood increased continuously from 8.1 to 16.3 and from 86 to 712, respectively.

The linker between the two domains and the termini were manually rebuilt in Coot [113] and the model was refined with BUSTER [119] using NCS restrains and two TLS groups per monomer in later stages of refinement. In the CZB domain, strong anomalous density was visible at the position, where a zinc ion is bound in $YdeH^{CZB}$. In the two active sites, well defined difference electron density was observed, which was comparable with GTPαS and a magnesium ion. Additionally, difference electron density was visible in the I-site of the GGDEF domains. In this density, part of a GTPαS molecule and a c-di-GMP molecule could be fitted. After further rounds of rebuilding in Coot and refinement with BUSTER using the model of the two individual domains as target restrains, the final model ended up in R and R_{free} values of 23.2 % and 23.4 %, respectively (Table 3.13).

Table 3.13: Refinement statistics for the YdeH full-length model.

Resolution[Å]	3.9
Space group	$P6_1$
R [%]	23.2
R_{free} [%]	23.4
Number of atoms	
protein	4349
nucleotides	158
magnesium	2
zinc	2
B factors	
main chain A / main chain B	68.1 / 52.8
side chain A / side chain B	78.4 / 62.1
nucleotides	62.7
magnesium	30.8
zinc	32.8
RMS deviations	
bond lengths [Å]	0.010
bond angles [°]	1.08
Ramachandran statistics [%]	
in favored regions	97.9
in allowed regions	1.3

Validation of the YdeH full-length model

Electron density was well defined for the whole main chain of full-length YdeH except for residues 36-53 of chain A, residues 37-64 of chain B as well as for the first four residues of chain A. Refinement progress was followed by monitoring the values of R and R_{free} and checking for model geometry. The final statistics are given in Table 3.13. The values for the R-factors were reasonable for the given data resolution of 3.9 Å and the RMS deviations of bond lengths and angles were in the standard range. Most of the main chain torsion angles were in the allowed regions of the Ramachandran plot, with most of them in the favored region (Figure 3.15).

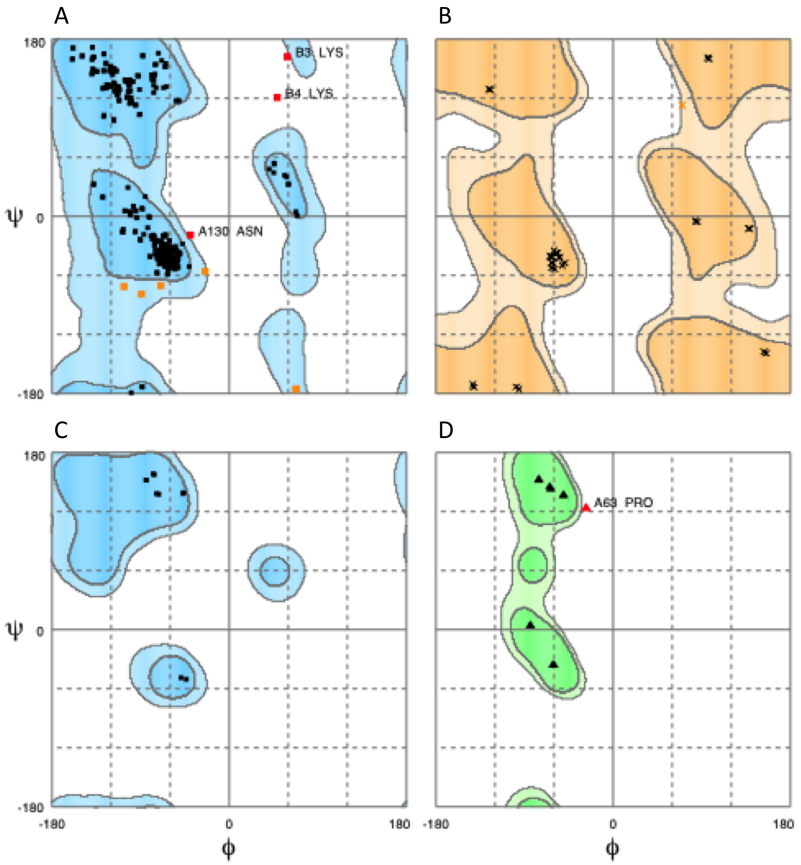
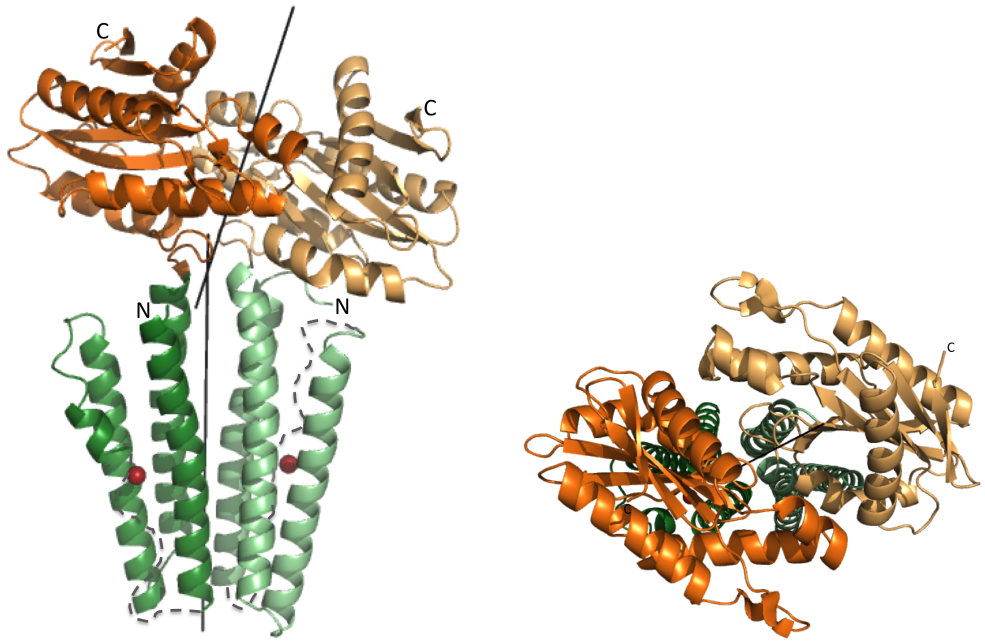


Figure 3.15: Ramachandran diagram for full-length YdeH. Main chain torsion angles were analyzed with RAMPAGE [116] and are indicated for each residue by crosses (glycines) or squares (all others). Dark and light background shading indicates energetically favorable and allowed regions, respectively. (A) General plot for all non-glycine residues. (B) Special plot for glycine residues. (C, D) Plots for pre-proline and proline residues, respectively.

3.1.5 Overall structure of YdeH

The structure of full-length YdeH (C52A mutant) was determined by molecular replacement to a resolution of 3.9 Å. YdeH folds into two domains, the N-terminal CZB domain and the C-terminal catalytic GGDEF domain. The overall shape of YdeH is that of a mushroom (Figure 3.16). The stipe is formed by the two CZB domains and the two GGDEF domains represent the cap that sits on the stipe. The CZB domains and the GGDEF domains are each related by a non-crystallographic two-fold axis. Since this axis is different for the two domains, the YdeH dimer is non-symmetric. This becomes apparent at the domain interface between the CZB and the GGDEF domain. While the domain contact in chain A is very small, the surface area in chain B is 4.5 times larger, being 550 $Å^2$. The presence of a dimer is in agrement with the results from gel filtration experiments (sec section 3.1.2).
The crystal contacts are listed in Table 3.14 and the crystal packing is shown in Figure 3.17. The largest crystal contact (II) with 625 $Å^2$ and another contact (III) is mediated via the 6_1-fold axis. The crystal contact IV via the 3_1-fold axis is additionally stabilized via a c-di-GMP and two GTPαS molecules. Thus, the two I_p-sites are connected with each other. The crystal solvent content of 57 % is not very high, but the formation of only three crystal contacts could be a reason for the low resolution of 3.9 Å.
The characteristics of the domains will be discussed in detail later. Next the structures of $YdeH^{CZB}$ and $YdeH^{GGDEF}$ will be analyzed. They were determined to a much better resolution of 2.2 Å and 1.8 Å, respectively.

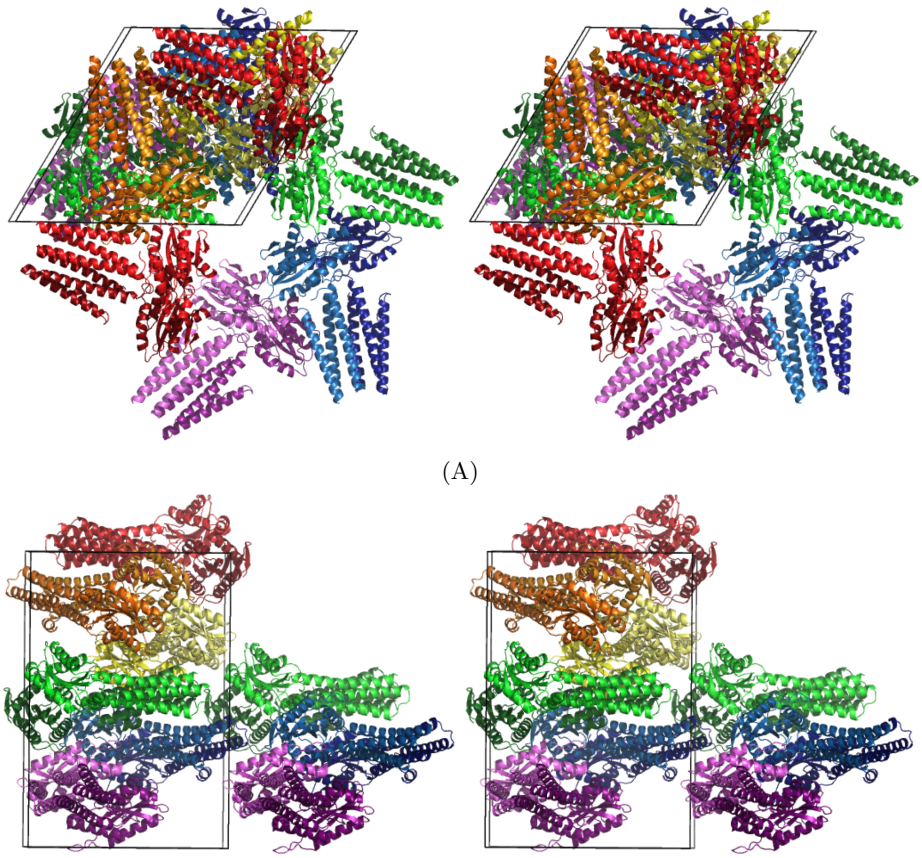


(A) (B)

Figure 3.16: Ribbon representation of YdeH. The CZB domains are colored in green and the GGDEF domains are colored in orange. Chains A and B are shown in dark and light colors, respectively. The bound zinc atom is depicted as sphere and colored in red. The disordered regions are marked as dashed lines. The non-crystallographic dimer axis for the CZB and GGDEF domain are indicated as black lines. (A) Side view. (B) Top view along the dimer axis of the CZB domain.

Table 3.14: Crystal contacts in the YdeH crystal.

Name	Contacting chains	Symmetry operation	Surface area [Å^2]	Interfacing residues chain 1	chain 2
I	A-B	NCS	2017.7	12, 13, 16, 19, 20, 22-24, 26-31, 33-35, 89, 90, 93, 98, 100, 103, 104, 106, 107, 110, 111, 113-115, 118, 121, 122, 125, 126, 129, 130, 133, 134, 140, 141, 144, 170, 173, 177-179, 182, 186, 204, 205, 207, 263, 266, 267, 270, 271, 274	12, 13, 16, 19, 20, 22-24, 26-31, 33-36, 89, 90, 93, 98, 100, 103, 104, 106, 107, 110, 111, 113-115, 118, 121, 122, 125, 128, 129, 130, 133, 134, 140, 141, 144, 170, 174, 177-179, 182, 185, 204, 205, 207, 209, 263, 266, 270
II	A-A	x-y-1, x-1, z+1/6	624.5	168, 171, 174, 175, 176, 227, 231, 234, 235, 244-246, 248, 280-283, 295	73, 74, 77, 78, 80, 81, 84, 87, 88, 91, 92, 97, 99, 101, 102, 104, 105, 108, 112
	B-A	x-y-1, x-1, z+1/6	10.9	271	99
III	B-B	x-y, x, z+1/6	195.4	152, 155, 258, 259	167, 168, 171, 246, 248, 280, 281, 282
IV	A-B	-y, x-y+1, z+1/3	102.2	60, 62, 63, 197, 239	17, 197, 239, 241, 242-244



(A)

(B)

Figure 3.17: Crystal packing of YdeH in space group $P6_1$ with one dimer (orange) per asymmetric unit. The content of the unit cell is shown and the coloring has been carried out by symmetry operations. Crystallographic related dimers are crosslinked by c-di-GMP (e.g. orange-green). (A) Stereo view along the z axis. (B) Stereo view along the y axis.

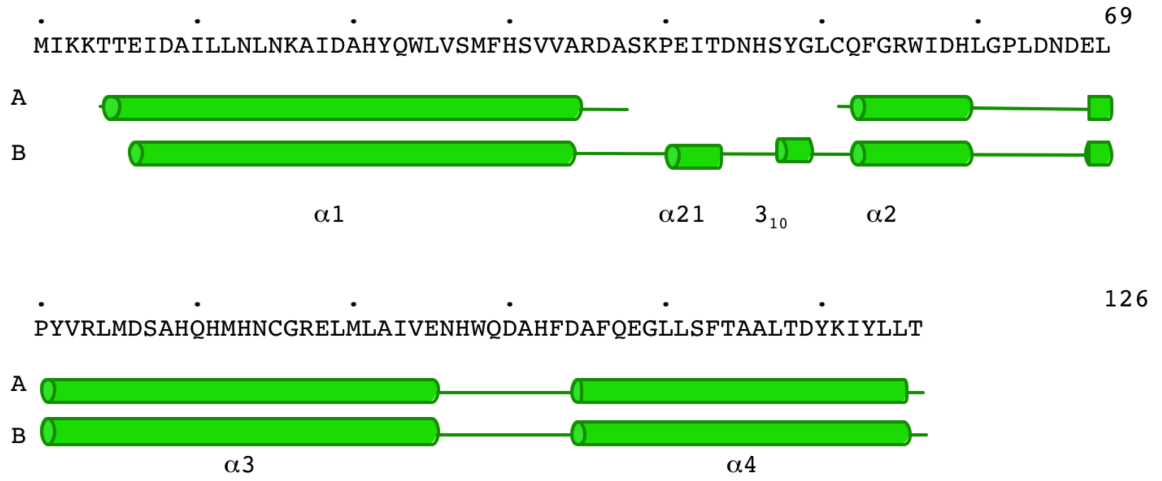


Figure 3.18: Secondary structure assignments for YdeH^{CZB} chain A and B. Every tenth residue is marked by a dot and disordered parts of the model are indicated by gaps.

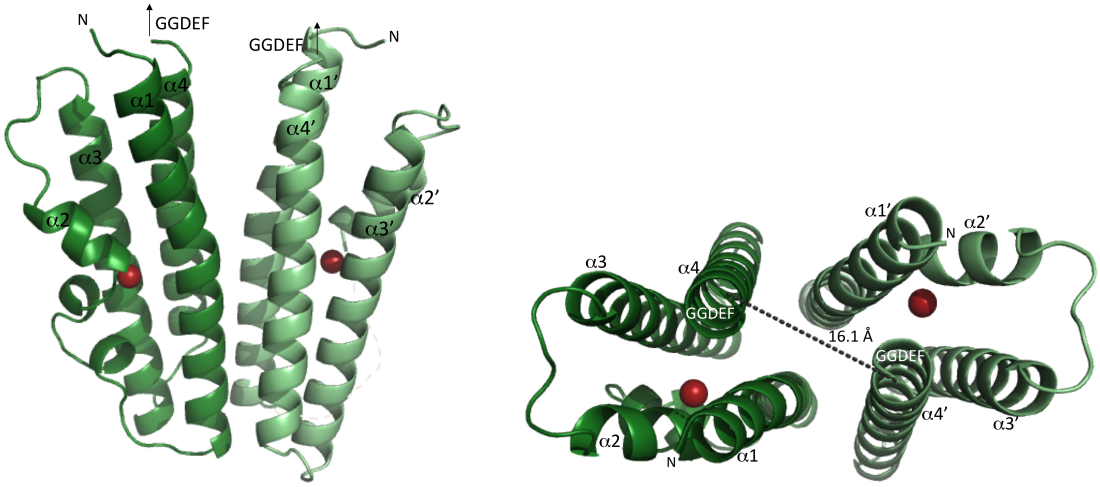
3.1.6 Structural analysis of the CZB domain

Structure of YdeH^{CZB}

The secondary structure of YdeH^{CZB} was assigned using DSSP [120] and STRIDE [121]. Both programs essentially delivered the same results, which is shown in Figure 3.18.

YdeH^{CZB} folds into a single elongated domain and forms a four helical bundle in which the helix $\alpha 2$ is rather short and is proceeded by a large loop (Figure 3.19). This part of the structure is not visible in chain A. The N- and C-termini are on the same side of the molecule and relatively close together in space. The distance between the C-termini of both chains is 16.1 Å. A zinc ion is located in the center of the helical bundle and mediates contacts between the four helices. The asymmetric unit consists of two monomers (Figure 3.19). This non-crystallographic dimer is thought to be functional *in vivo*. From size exclusion experiments (see section 3.1.2) it is known that YdeH^{CZB} forms dimers in solution. The dimer interface is formed by helices $\alpha 1$ and $\alpha 4$. The dimer contact covers 1278 Å^2, which corresponds to 17 % of the total surface area of a monomer and is mediated by ten H-bonds, two salt bridges, four aromatic stackings and several non polar interactions (Figure 3.20).

In addition to the dimer interface, there are two dominant crystal contacts (II and III) present. The interactions lie along the y-axis. Crystal contacts are listed in Table 3.15 and the crystal packing is shown in Figure 3.21. The YdeH^{CZB} dimer makes seven contacts with other molecules to build up the lattice. The total surface area of the dimer involved in crystal contacts is 1560 Å^2 or 12.2 % of the complete surface area. While the residues 38-51 of chain B are involved in crystal contacts (contact IV) and chain A is not, the unstructured loop of chain A has more freedom and is not visible in the electron density (Figure 3.22).



(A) (B)

Figure 3.19: Ribbon representation of a YdeH^{CZB} dimer. Chains A and B are colored in green and dark green, respectively. The bound zinc atom is depicted as a sphere and colored in red. The disordered region in monomer A is marked as a dashed line. The C-terminal GGDEF domain would be located on the top. The distance between the C-termini (Tyr126) is 16.1 Å. (A) Side view. (B) View along the dimer axis.

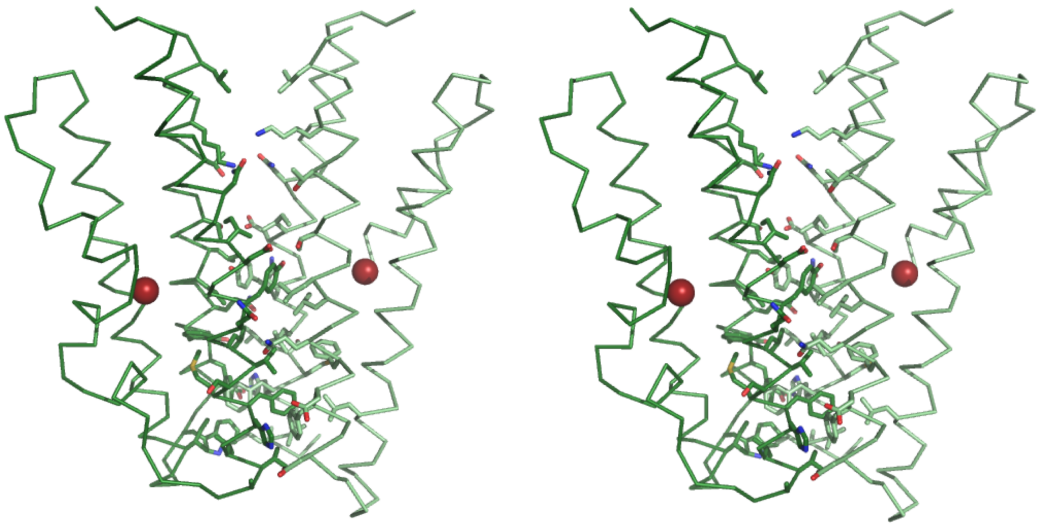


Figure 3.20: Stereo view of the dimer interface of YdeH^{CZB}. The protein backbone of chain A and B is shown in green and dark green, respectively. The bound zinc ion is shown as a red sphere and the interacting residues at the dimer interface are depicted as sticks. The dimer interface is formed by helices α1 and α4 and is mediated via H-bonds and salt bridges, aromatic stackings and non polar interactions.

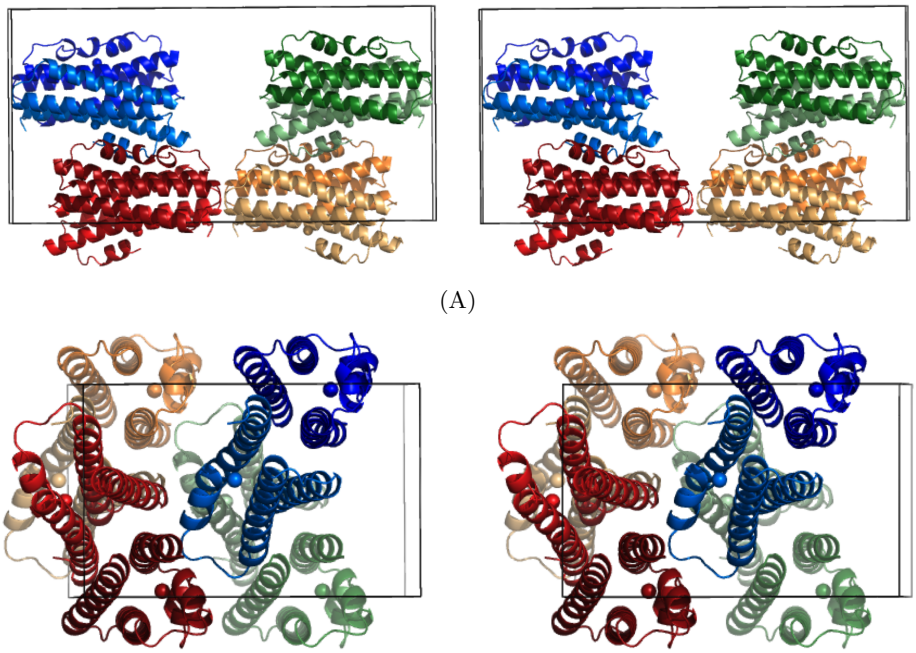


Figure 3.21: Crystal packing of YdeH^{CZB} in space group C2 with one dimer (green) per asymmetric unit. The content of the unit cell is shown and the coloring has been carried out by symmetry operations. (A) Stereo view along the z axis. (B) Stereo view along the x axis.

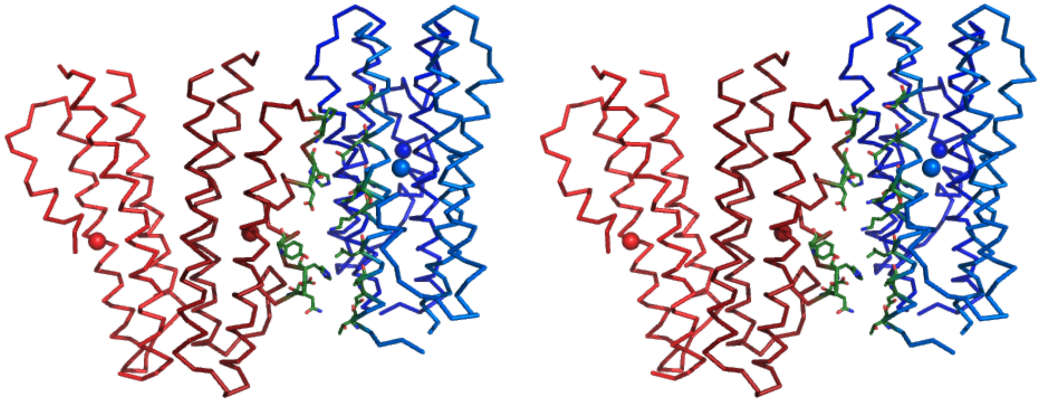


Figure 3.22: Stereo view of the crystal contact IV of the YdeH^{CZB} crystal. Monomer A is shown in light and monomer B in dark colors. The interfacing residues are shown as green sticks. The large loop of monomer B (dark red) is involved in crystal contacts and has no freedom to adopt other conformations.

Table 3.15: Crystal contacts in the YdeH^{CZB} crystal.

Name	Contacting chains	Symmetry operation	Surface area [Å^2]	Interfacing residues chain 1	chain 2
I	A-B	NCS	1277.3	16, 19, 20, 22-24, 26-31, 33-35, 89, 90, 93, 98, 100, 103, 104, 106, 107, 110, 111, 113, 114, 118, 121, 125	16, 19, 20, 22-24, 26-31, 33-35, 89, 90, 93, 98, 100, 103, 104, 106, 107, 110, 111, 113, 114, 118, 121, 125
II	B-A	x, y, z-1	364.2	66, 69-71, 73, 74, 77-78, 81, 84, 88, 115, 119, 122	66, 67, 69-71, 73, 74, 77, 78, 81, 84, 88, 119
III	A-B	-x+1/2, y-1/2, -z+1	351.8	59, 60, 61, 62, 63-66, 68, 73, 76	17, 20, 21, 24, 25, 28, 40-43, 51, 53, 56
	A-A	-x+1/2, y-1/2, -z+1	95.6	6, 7, 60-62	100, 101, 104, 107, 111
IV	A-B	-x+1/2, y-1/2, -z	202.9	6, 9, 10, 13, 17, 20, 24, 28, 53	45-47, 49, 59, 60, 62, 63
	B-B	-x+1/2, y-1/2, -z	66.2	107, 111	60-62
V	B-B	-x, y, -z	180.9	33, 34, 36, 94-96	33, 34, 36, 94-96
VI	B-A	-x, y, z	162.6	95, 96, 97, 98, 99	36, 93, 94, 96, 98
VII	A-A	-x, y, -z+1	136.3	95-97	95-97

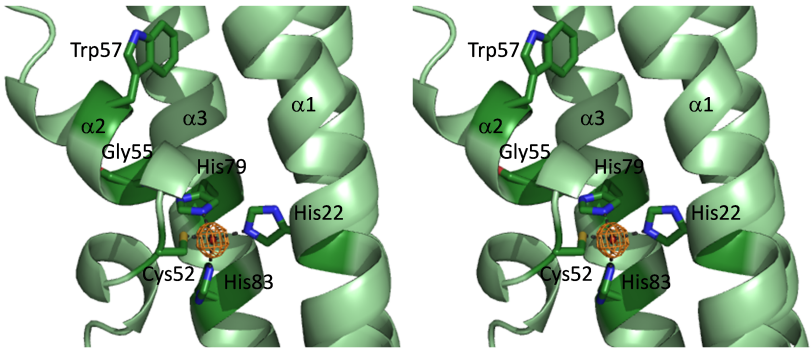
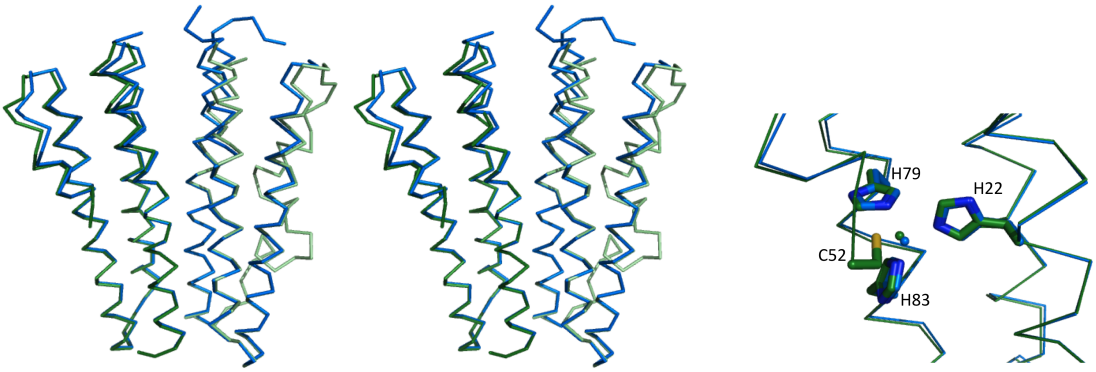


Figure 3.23: Stereo view of the zinc binding site of YdeH^{CZB}. The zinc coordinating residues (His22, Cys52, His79, His83) and the conserved residues Gly55 and Trp57 are shown as sticks. The zinc ion is depicted as a red sphere. The omit map for the zinc is countered at 4 σ.

Table 3.16: Zinc coordination geometry in YdeH^{CZB}.

Residue	Distance to zinc in chain A / chain B [Å]
His22	2.23 / 2.11
Cys52	2.34 / 2.32
His79	2.20 / 2.16
His83	2.15 / 2.17



(A) (B)

Figure 3.24: Superposition of the Cα-backbone of the CZB domain of YdeH^{CZB} (green) and full-length YdeH (blue). (A) Total stereo view and (B) zoom in the zinc binding site of monomer A.

Zinc binding

An unexpected finding during structure solution of YdeH^{CZB} was the presence of a zinc ion. This was surprising, as zinc had not been added during protein purification or crystallization. A first indication for the presence of zinc was already seen in the fluorescence scan during data collection and also ShelX found two more heavy metal sites during phasing. The zinc is coordinated by His79 and His82 from helix α3, His22 from helix α1 and Cys52 from the beginning of helix α2. The coordination is tetrahedral with ideal distances comparable to other zinc coordinations [137], and is virtually identical in both chains (Table 3.16). A detailed view of protein zinc interactions is given in Figure 3.23.

A zinc binding site consisting of three histidines and one cysteine is exceedingly rare. A search of the PDB with PDBeMotif [126] for structures containing zinc binding sites revealed over 6850 structures, but only ten with a 3His1Cys arrangement as in YdeH. These include five members of the matrix metalloproteinase family, the human Stromelysin-1 (PDB: 1slm), the human matrix metalloproteinase-2 (PDB: 1ck7), the human matrix metalloproteinase MMP-9 (PDB: 1l6j), the human matrix metalloproteinase MMP-1 (PDB: 1su3) and the human morphogenetic protein 1/tolloid-like metalloproteinase (PDB: 3edi). These proteins

are synthesized as inactive precursors and contain a catalytic zinc ion that is coordinated by three histidines from the catalytic domain and one cysteine from the N-terminal propeptide. The metalloproteinases get activated by interruption of the Cys-Zn interaction (cysteine-switch mechanism) and as a result, the active site is now accessible for the substrate and the fourth coordination partner of the zinc is available for a water molecule [138–142].
The second group of enzymes containing the 3His1Cys coordination are the class II aminoacyl-tRNA synthetases. In the editing domain of threonyl-tRNA synthetase from *Staphylococcus aureus* (PDB: 1nyr), of alanyl-tRNA synthetase from *Pyrococcus horikoshii* OT3 (PDB: 1v4p) and of alanyl-tRNA synthetase from *Archaeoglobus fulgidus* (PDB: 2ztg), the zinc ion bridges the two subdomains and might activate the carbonyl group of a mismatching amino acid in the deacylation catalysis [143–145].
Another protein with this zinc coordination is the human AMSH-LP deubiquinating enzyme (PDB: 2znv). The function of the zinc coordination is the accurate positioning of the substrate to allow efficient deubiquitination [146].
In the zinc-binding domain of the viral envelope glycoprotein from Junin virus (PDB: 2loz), the 3His1Cys zinc coordination has a structural function to maintain the correct fold [147].
An additional hit was the engineered human carbonic anhydrase II (PDB: 1dca). In this structure, an additional cysteine was introduced to the naturally occurring 3His-zinc binding motif, to inactivate the enzyme by replacing the catalytically important water molecule [148].
The zinc binding site consisting of three histidines and one cysteine in these proteins adopt different functions *in vivo*. A pure structural role, protein activation by ligand exchange and a catalytic function can be performed by using this zinc binding site.

In YdeH, the zinc is coordinated by four protein ligands and can therefore be assigned as a structural zinc site. On the other hand, after removal of zinc by incubation with a chelator, YdeH was still stable and well folded as evidenced by gel filtration experiments (section 3.1.10). This suggests that zinc does not only display a pure structural role. YdeH might be able to sense the zinc concentration as it is the case for zinc-responsive signal transduction proteins and in this way could regulate YdeH activity. The unstructured loop right before the Cys52 is ideally suited to allow fast exchange kinetics. Furthermore, the regulation of the GGDEF domain can be performed by a redox switch of the Cys52. The affinity is reduced by oxidizing the cysteine so that the zinc can not bind anymore. Kinetic experiment were performed as described in section 3.1.10 to elucidate the specific role of the zinc.

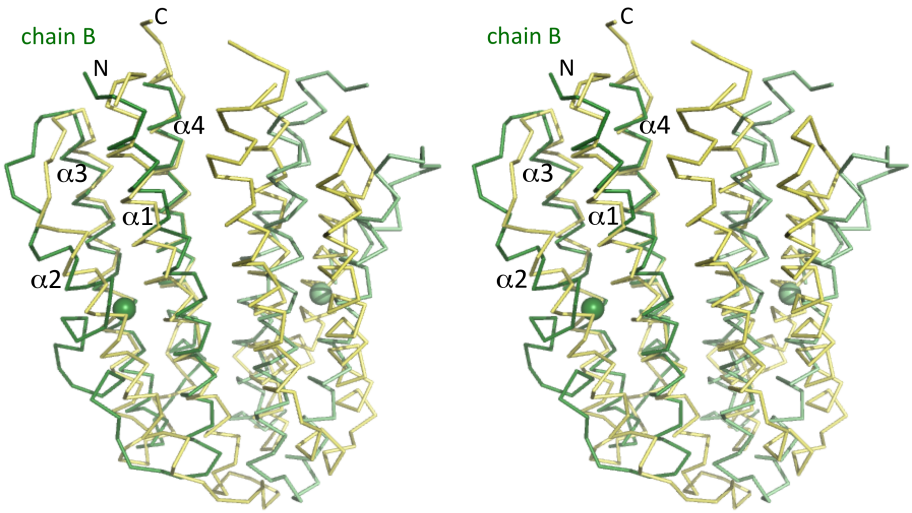


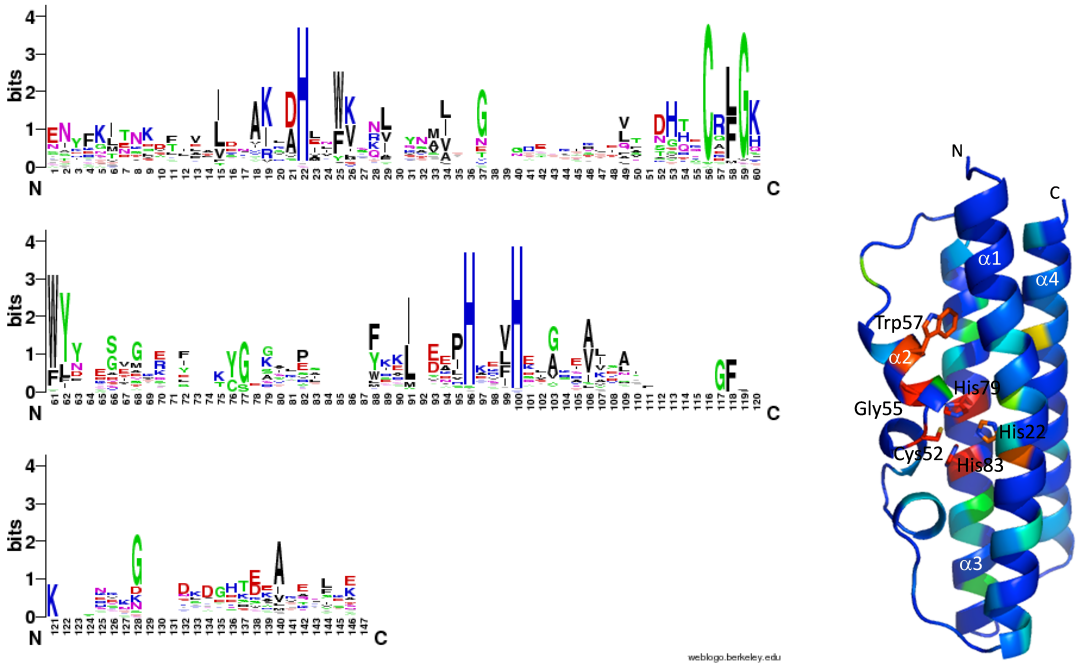
Figure 3.25: Stereo view of the superposition of YdeH^{CZB} chain B (green) with the ligand-binding domain of the Tar receptor from *S. typhimurium* (yellow). The helices α1, α3 and α4 superimpose quite well, but there are differences in helix α2 and the dimer orientation towards each other.

Structural comparison of YdeH^{CZB} and the CZB domain of full-length YdeH

The folding of the CZB domain is the same in the YdeH^{CZB} and full-length YdeH structure. A superposition of both models is shown in Figure 3.24. The RMS deviation of the Cα position is 0.7 Å and there are only minor differences at the termini. The dimer interface is identical. One difference is that in the full-length structure helix α2 and the preceding loop region is not visible in one subunit. Furthermore, the mutant C52A was used for determination of the YdeH full-length structure. This mutation does not impede the zinc binding to YdeH, but makes the zinc binding site more accessible so that the unstructured loop can adopt a completely free conformation, which is consequently disordered as evidenced by missing electron density.

Comparison with homologous proteins

To assess structural similarities between YdeH^{CZB} other proteins, a search for structural homologous was performed. The Dali program found only one hit with a Z-score above ten (10.2), the ligand-binding domain of the aspartate receptor Tar from *Salmonella typhimurium* (PDB: 1vlt). Monomer B of YdeH^{CZB} has a RMS deviation of 1.6 Å to monomer B of the aspartate receptor and a sequence identity of 14 %.
The Tar receptor is an inner membrane spanning chemosensory protein that interacts with the histidine kinase CheA. Upon binding of one of its cognate ligands (aspartic acid or maltose-binding protein), Tar modulates the autophosphorylation activity of CheA, which

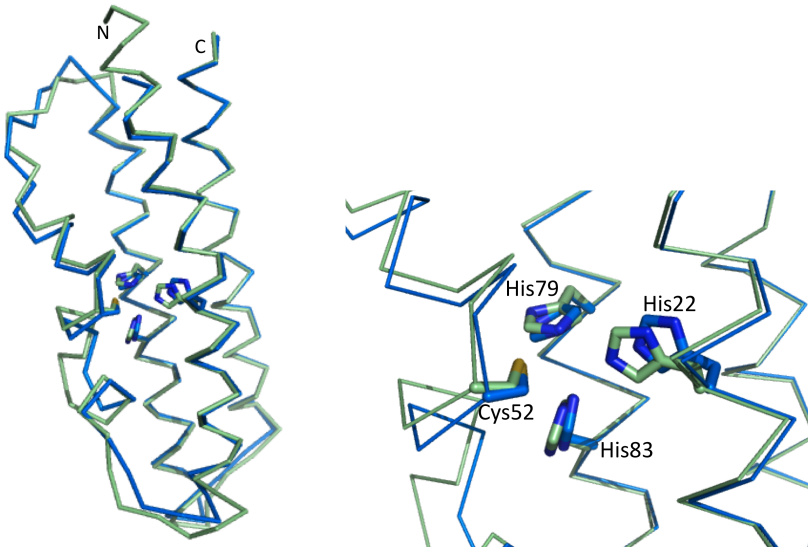


(A) (B)

Figure 3.26: Structure based sequence alignment of the members of the Pfam domain PB0001058 with the CZB domain of YdeH. The sequence alignment was done with the programm ALAdeGAP [127]. (A) Sequence alignment shown as WebLogo [128]. The letter height represent the degree of conservation. (B) Conservation mapping of the sequence alignment performed with the program protskin [129]. The color ranges from blue (not conserved) to red (strongly conserved).

in turn regulates the swimming behavior of the bacterium. Within the TAR family, the only characteristic is the overall conservation pattern of hydrophobic and polar residues that reflects the coiled coil nature of this domain. There are no obvious conserved binding sites, which is expected as known homologs bind very different ligands [149].
Both proteins, YdeH^{CZB} and Tar, fold into a four helical bundle and are able to form dimers. However, the dimer orientation towards each other are different (Figure 3.25). The most outstanding difference is the absence of the zinc binding motif in the aspartate receptor and differences in helix α2. While the aspartate receptor forms a long α2 helix, YdeH^{CZB} has a short one preceded by a unstructured loop region. The structural similarity of the N-terminal domain of YdeH with the sensory domain of the Tar receptor is based on the characteristics of a four helical bundle with hydrophobic residues pointing to the inner of the protein and polar residues pointing outside. It seems that this four helical bundle fold is suited to transfer signals to associated domains.

As no other structures of CZB proteins are known, a structure based sequence alignment



(A) (B)

Figure 3.27: Superposition of the model of TlpD from *H. pylori* (blue) on chain B of YdeH^{CZB} (green). The model of TlpD was generated with MODELLER based on a structural sequence alignment with YdeH^{CZB}. (A) Ribbon representation. The zinc binding residues are shown as sticks. (B) Close up view of the zinc binding site.

with members of the new Pfam domain PB0001058, where YdeH^{CZB} belongs to, was performed. The alignment carried out with WebLogo and a mapping of conserved residues are shown in Figure 3.26. A high degree of conservation is identified for the zinc binding motif (His22, His78, His83, Cys52) and for a region of helix α2 containing Gly55 and Trp57. The conservation of Gly55 is probably due to steric reasons to allow a correct zinc coordination. The role of the aromatic residue Trp57 is not really obvious. The side chain of Trp57 is not completely hidden in the structure, but is also not completely solvent exposed. This region may be a possible interaction area for other proteins to give a input signal that has to be transferred to the catalytic output domain. The transfer of the signal could be via a zinc release. The hydrophobic core of the helical bundle is conserved, that is not the case for the dimer interface. For this domain the zinc binding seems to be a crucial function, but the dimerization of the CZB domain seems to be quite unusual.

The only characterized member of this new Pfam domain is the cytosolic chemoreceptor TlpD from *H. pylori*. In addition to the CZB domain, TlpD contains a methyl-accepting chemotaxis-like (MA) domain. A model of TlpD was created with MODELLER [130] based on the structural sequence alignment with YdeH^{CZB} performed with ALAdeGAP [127] (Figure 3.27). The model displays a RMS deviation of the Cα positions of 1.2 Å to YdeH^{CZB} and has the same arrangement of the zinc coordinating residues. This is in agrement with the experimental results of TlpD, where zinc binding was shown by ICP-MS experiments [52]. It

seems that in this domain, which often occurs in combination with signaling domains, zinc acts as the sensory signal, which transfers the signal to the associated domain to modulate its activity.

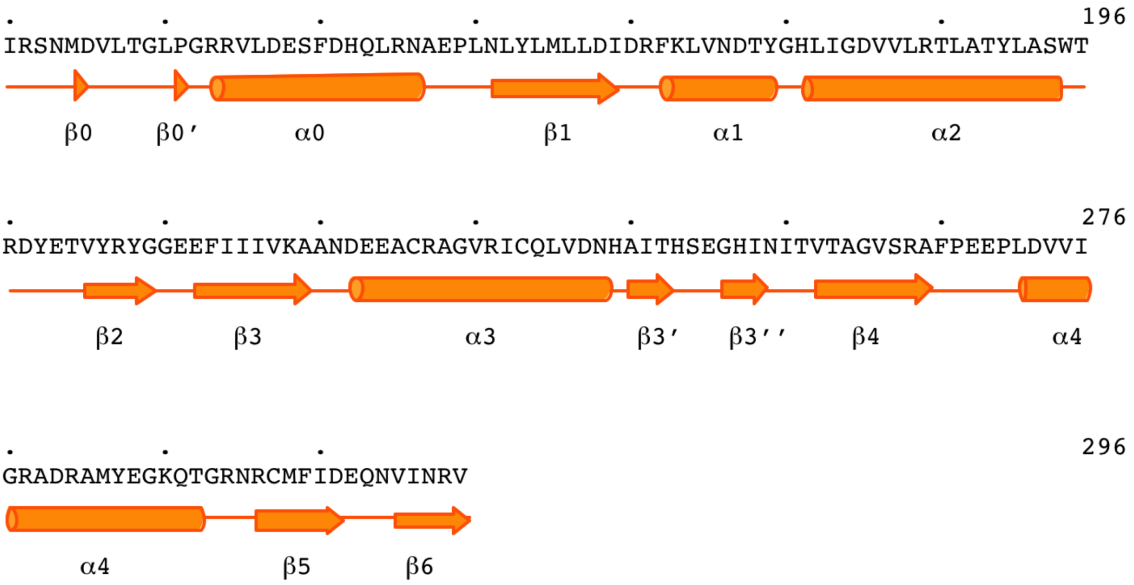


Figure 3.28: Secondary structure assignments for YdeH^{GGDEF}. Only one chain is shown, because the two monomers have the same secondary structure assignments. Every tenth residue is marked by a dot.

3.1.7 Structural analysis of the DGC domain

Structure description of YdeH^{GGDEF}

Secondary structure elements of the YdeH^{GGDEF} structure were determined with DSSP[120] and STRIDE [121] with an identical result, which is shown in Figure 3.28. The DGC domain of YdeH consists of five α-helices and eight β-strands.
YdeH^{GGDEF} folds into a globular domain with a diameter of 46 Å and exhibits the typical fold of diguanylate cyclases. These domains consist of a central five-stranded β-sheet surrounded by helices. Compared to the structurally best known DGCs PleD and WspR, YdeH has an additional β-strand at the C-terminus. The topology of the GGDEF domain of YdeH is depicted in Figure 3.29 and the overall fold is shown in Figure 3.30. The active site with the signature motif GGDEF is located on the loop between the β-strands β2 and β3. The I_p-site is located on the opposite side of the protein between helix α2 and β-strand β2.

YdeH^{GGDEF} crystallized with one monomer in the asymmetric unit, which is also the major species in solution, known from gel filtration experiments (section 3.1.2). YdeH must form as a dimer to be active. Dimerization is probably mediated via the CZB domain and not the GGDEF domain, since the GGDEF domain requires some conformational flexibility during the catalysis. Indeed, all existing crystal contacts are between 380 and 114 Å^2 and are

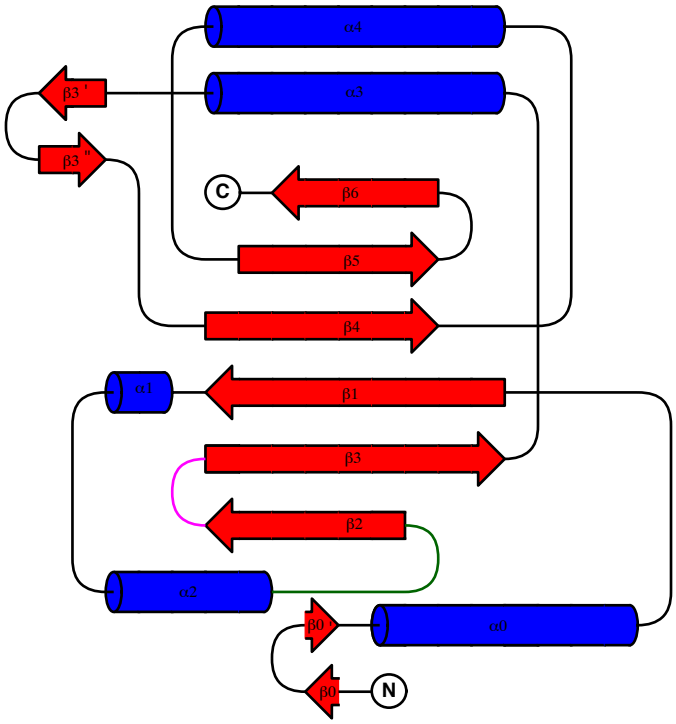
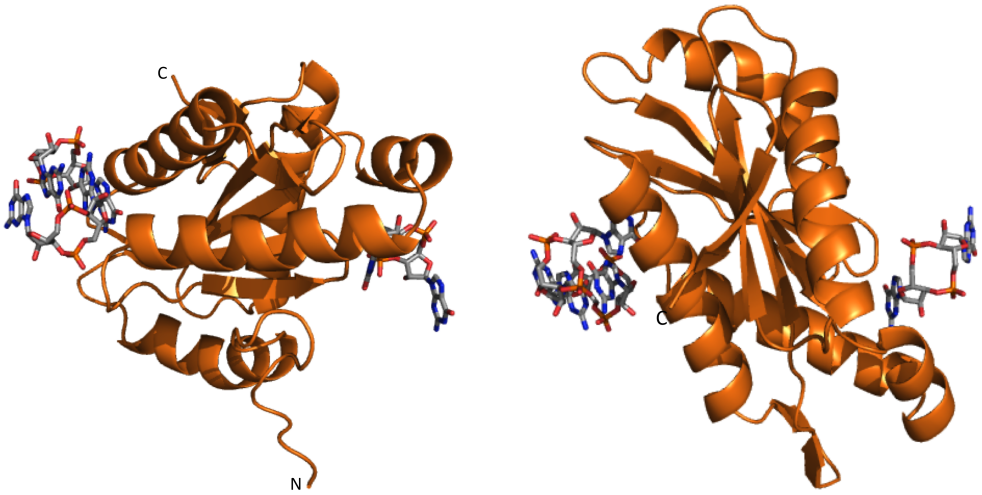


Figure 3.29: Topology of the GGDEF domain of YdeH. α-helices are depicted as blue cylinders and β-strands as red arrows. The active side loop is marked in magenta and the I-site loop in green. The plot was created with TopDraw [150].



(A) (B)

Figure 3.30: Ribbon representation of the GGDEF domain of YdeH. C-di-GMP bound to the active site and a c-di-GMP dimer bound to the I_p-site are shown as sticks. (A) Side view and (B) top view.

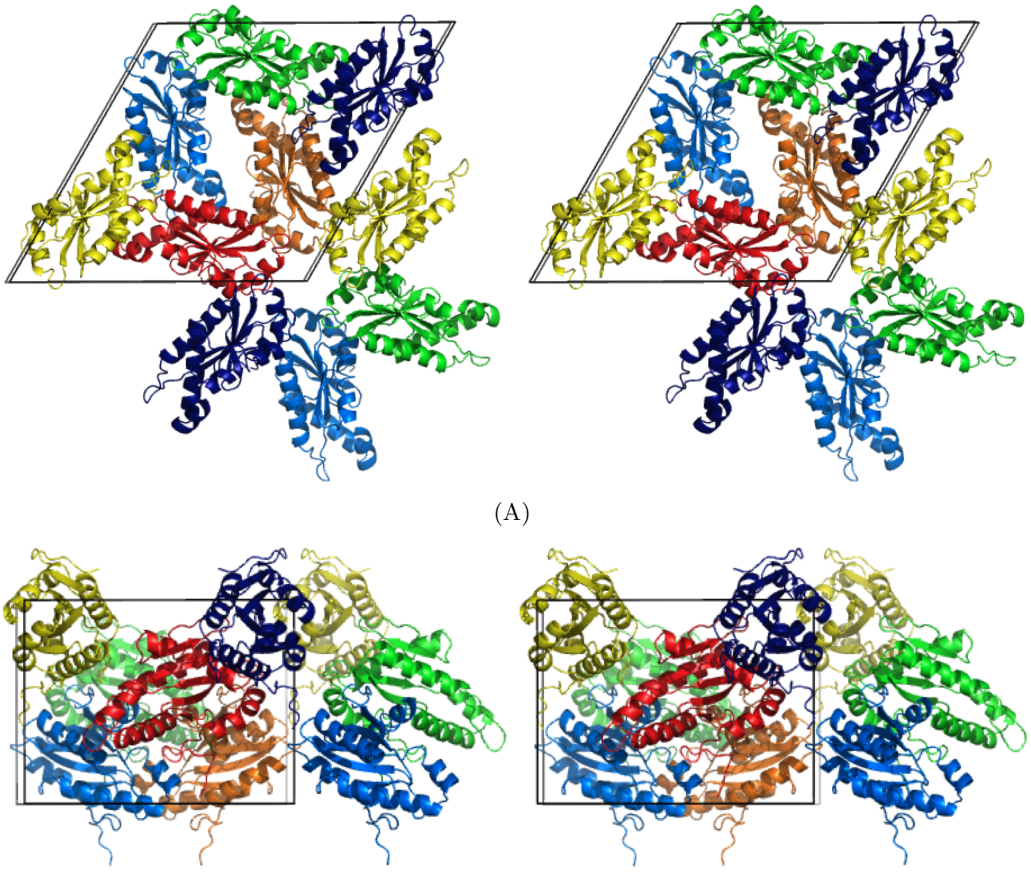
Table 3.17: Crystal contacts in the YdeH^{GGDEF} crystal.

Name	Symmetry operation	Surface area [Å^2]	H-bonds/ salt bridges	Interfacing residues chain 1	chain 2
I	x-y, x, z+1/3	379.8	0/0	149, 156, 157, 197-199, 215, 217	258-260, 263, 264, 267, 268, 271, 287, 289-291
II	x-y, x, z-2/3	192.3	4/1	217, 218, 220, 221, 224	126, 127, 128, 129, 141
III	-y+1, x-y+1, z-1/3	191.5	1/0	168, 171, 231, 234, 235, 279, 280, 282, 283, 295	191, 194, 195, 235, 236, 239, 240, 241, 243, 244
IV	-y+1, x-y+1, z+2/3	127.5	0/0	179, 180, 241, 242	283, 295, 297
V	-x+1, -y+1, z	114.2	0/2	140, 170, 174, 209, 274	140, 170, 174, 209, 274
VI	x, y, z-1	27.5	0/0	290, 292	126, 128

therefore to small to form a dimer in a biological assembly. All crystal contacts are listed in Table 3.17 and the packing is shown in Figure 3.31. In contact V, the two GGDEF domains are linked over the two-fold axis via a c-di-GMP molecule in the two active sites (see Figure 3.34). The surface area is only 114 Å^2 but an additional contact area of 235 Å^2 is generated by c-di-GMP. This interaction is stabilized by two salt bridges. Such an arrangement is thought to be the end point of catalysis, where the c-di-GMP has already formed but has not left the active site, yet. Furthermore, two GGDEF monomers are cross linked with a c-di-GMP dimer bound to the I-site via the six-fold axis (contact I) (see Figure 3.37). This arrangement might represent a natural inhibition mode.

Structure description of the GGDEF domain in full-length YdeH

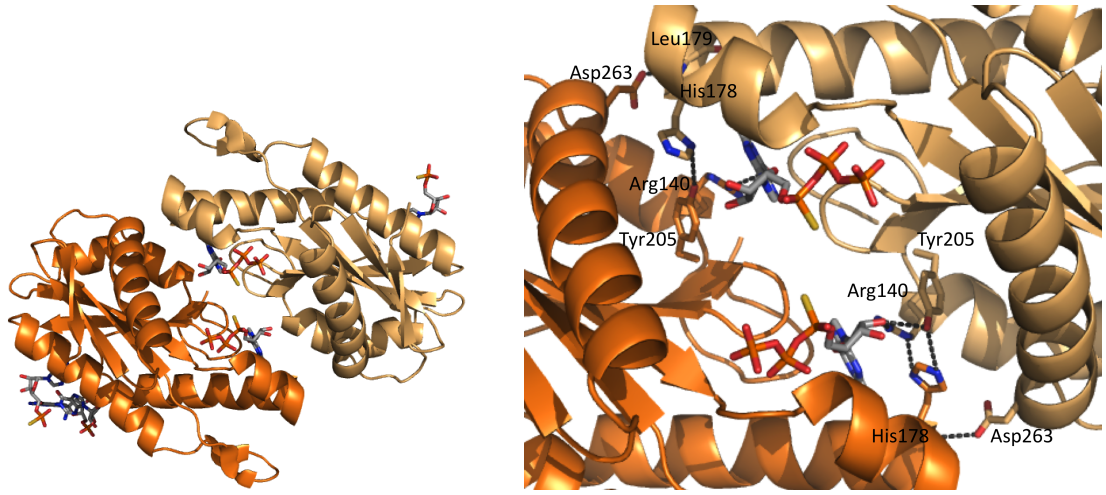
Full-length YdeH crystallized with two monomers in the asymmetric unit and had the substrate analog GTPαS bound to both active sites. Furthermore, a c-di-GMP with a stacked GTPαS was bound to the I-site. The structure is shown in Figure 3.32. The dimer interface between the two monomers in the asymmetric unit covers an area of 719 Å^2, whereof 536 Å^2 is made by protein-protein interaction and the remaining 183 Å^2 are contacts between protein and GTPαS (Figure 3.32). The contact is mediated via five H-bonds between the proteins and three between protein and substrate analog (Table 3.18). The dimer interface is larger than in the YdeH^{GGDEF} structure but is still not large enough to form a biological assembly. One reason for this could be that the catalysis takes place at this interface and



(A)

(B)

Figure 3.31: Crystal packing of YdeH^{GGDEF} in space group P6$_2$ with one monomer (orange) per asymmetric unit. The content of the unit cell is shown and the coloring has been carried out by symmetry operations. C-di-GMP that crosslinks two active sites lies on the two-fold axis in the middle of the cell and crosslinks the orange and light-blue molecule (contact V). The linkage of c-di-GMP of the I-sites is via the six-fold axis for example between the orange and the red molecule (Contact I). (A) Stereo view along the z axis. (B) Stereo view along the y axis.



(A) (B)

Figure 3.32: Structure of the GGDEF domain in full-length YdeH. (A) Ribbon representation of the dimer. Chain A and B are colored in dark and light orange, respectively. The GTPαS and c-di-GMP molecules are shown as sticks. The view is along the two-fold axis. (B) Close up view of the dimer interface. The interacting residues are shown in stick representation.

therefore the interface must be rather variable to allow product binding, catalysis and substrate release.

The overall fold of full-length YdeH is identical to the YdeH^{GGDEF} model with Cα RMS deviation of 0.12 Å (Figure 3.33).

Product binding to the active site

The active site of the DGCs is defined by the GGDEF motif, located on the turn between the β-strands β2 and β3. In the YdeH^{GGDEF} structure, the reaction product c-di-GMP is bound in the active site and thereby symmetrically crosslinks two GGDEF domains (Figure

Table 3.18: H-bonding in the dimer interface of the GGDEF domains of full-length YdeH.

Chain A	Chain B	Distance [Å]
His178-$N_{\delta 1}$	Arg140-N_{ϵ}	2.9
His178-$N_{\epsilon 2}$	Tyr205-OH	3.3
His178-N	Asp263-$O_{\delta 2}$	3.0
Tyr205-OH	His178-$N_{\delta 1}$	3.3
Asp263-$O_{\delta 2}$	Leu179-N	3.0
Arg140-$N_{\eta 2}$	GTPαS-O6	2.9
GTPαS-O2'	Arg104-$N_{\eta 2}$	2.9
GTPαS-O2'	Tyr205-OH	2.7

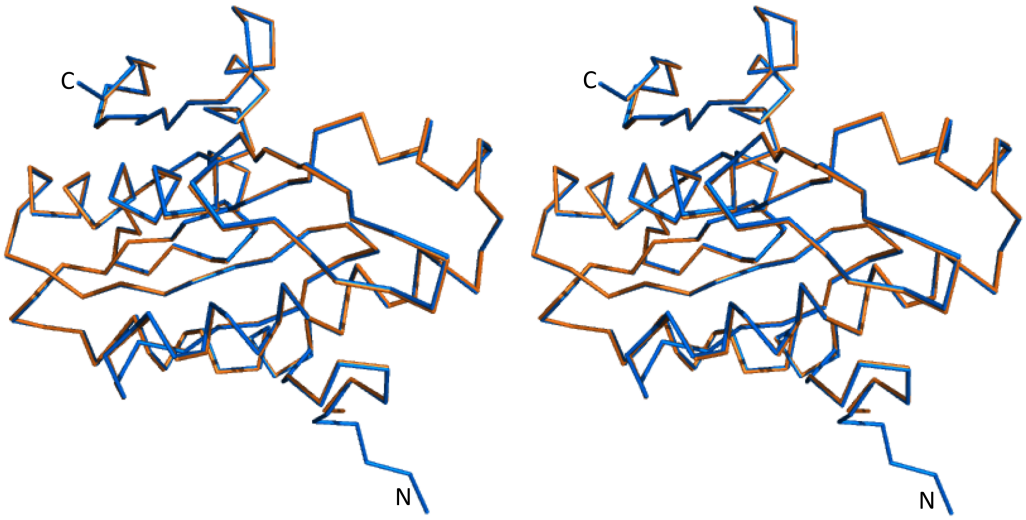
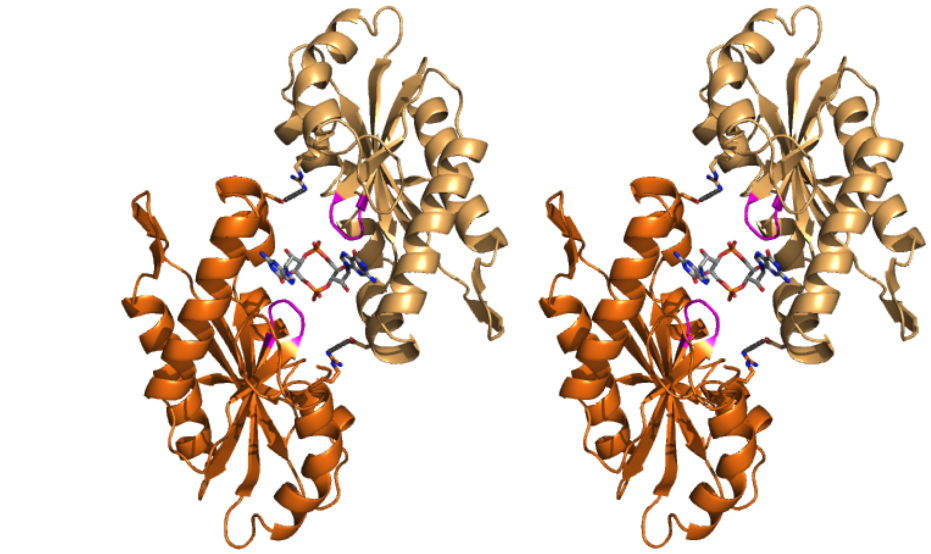


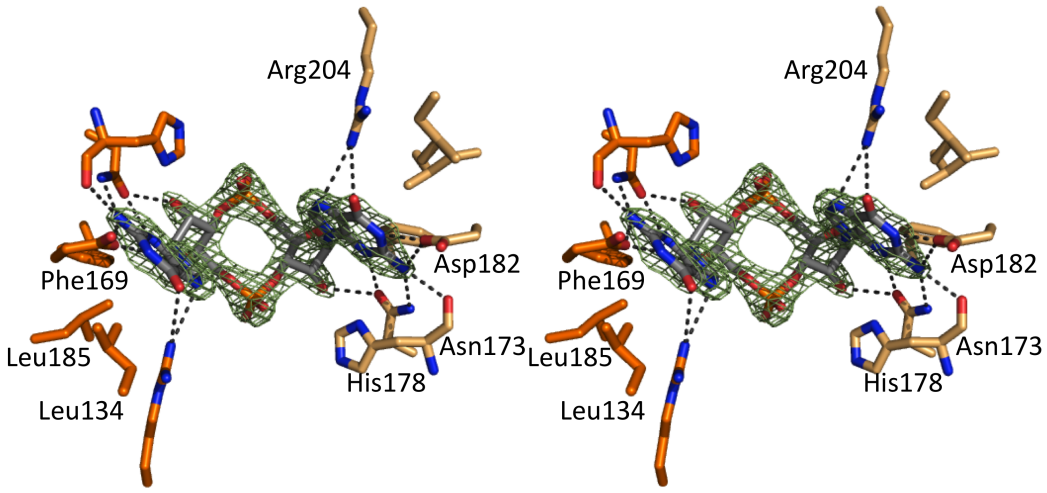
Figure 3.33: Stereo view of the superposition of the GGDEF domain of YdeH^{GGDEF} (blue) and full-length YdeH (orange).

3.34). The guanyl base of c-di-GMP is fixed in a hydrophobic pocket, which is formed by Leu134, Phe169 and Leu185. In addition, the base is positioned by H-bonding with Asn173, Asp182, Arg204 and the backbone oxygen of His178. Asn173 is interacting with the 2'OH of the ribose. The phosphate groups are not in contact with the protein. A detailed view of the active site is given in Figure 3.34 and the distances for the interacting residues are listed in Table 3.19.
The present binding arrangement could reflect the end point of catalysis, where the reaction product c-di-GMP has already been formed, but has not left the active site, yet. The dimeric GGDEF arrangement may be stabilized by two salt bridges between the conserved residues Arg140 and Asp174. The N-termini (Ile127) of the GGDEF domains are located next to each other with a distance of 26 Å, which is 9 Å more than the distance of the C-termini in the CZB domain in the YdeH^{CZB} structure.

So far, two other DGC structures are known, which have c-di-GMP bound in the active site. In the non-activated structure of PleD (PDB: 1w25) [15], c-di-GMP is present in the active site and thereby crosslinks two symmetry related GGDEF domains, which are from different dimers (Figure 3.35). The binding of c-di-GMP is similar as in YdeH^{GGDEF}. The guanyl base is positioned in the same hydrophobic pocket and the interacting residues are the same, apart from Gly369. The homologous residue Gly207 seems not to be in direct contact in YdeH^{GGDEF}. The dimer arrangement of the GGDEF domains in both structures is slightly different as shown in Figure 3.35. Therefore, the c-di-GMP can not be aligned perfectly.



(A)



(B)

Figure 3.34: The active site of YdeH^{GGDEF}. (A) Ribbon representation of the dimer, which is crosslinked by c-di-GMP in the two active sites. The active site loop is marked in magenta and c-di-GMP and the amino acids involved in the salt bridge formation (Arg140, Asp174) are shown as sticks. The view is along the two-fold axis from the N-terminus. (B) Stereo view of the c-di-GMP coordination geometry at the active site. The interacting protein residues and c-di-GMP are shown in stick representation. The omit map for c-di-GMP is countered at 2.5 σ.

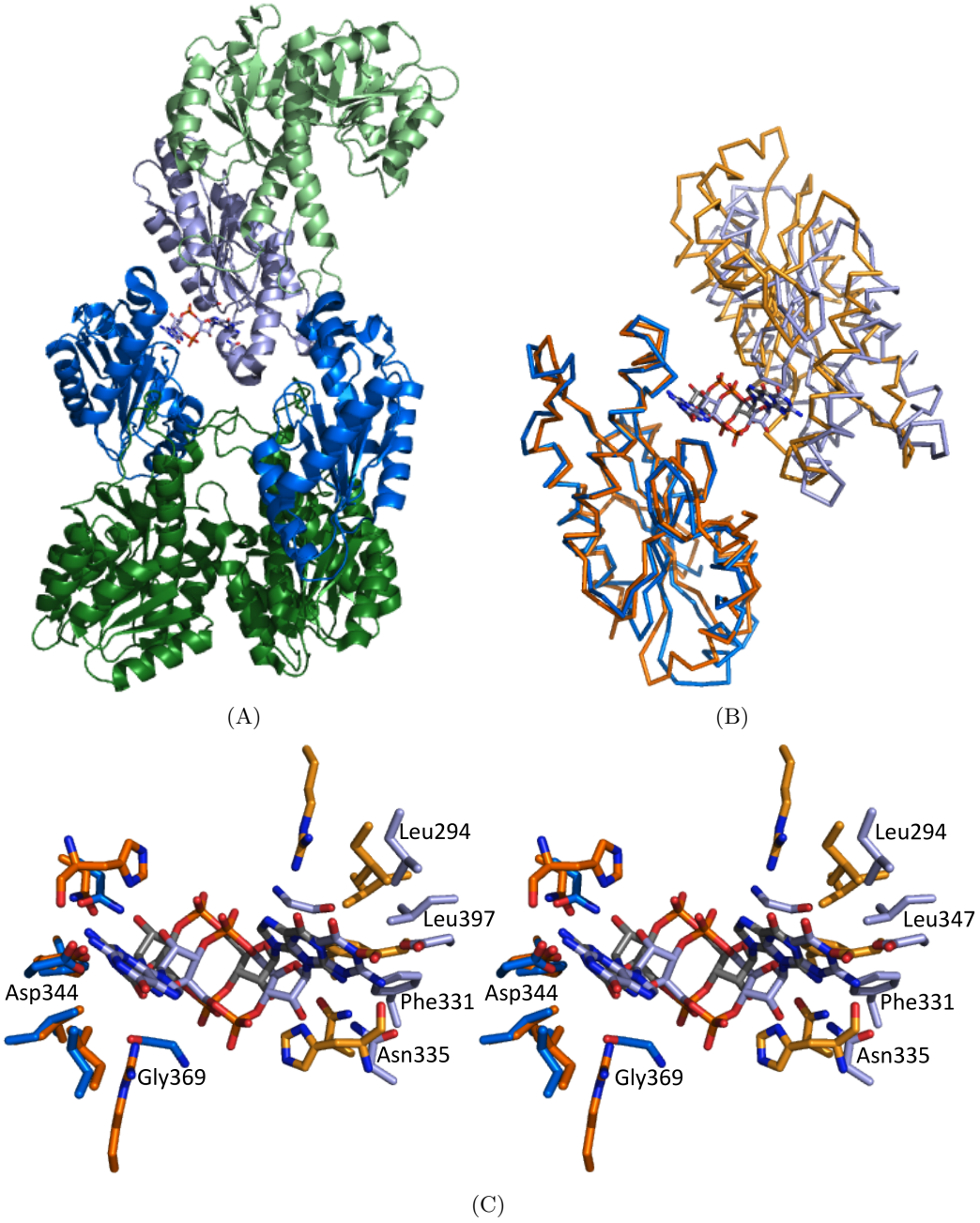


Figure 3.35: C-di-GMP binding to the active site of PleD (PDB: 1w25) compared to YdeH^{GGDEF}. (A) Ribbon representation of non-activated PleD. The DGC domain of the dimer is shown in blue and the REC domains are shown in green. A symmetry related monomer is shown in light colors, which is crosslinked via c-di-GMP in the active site. C-di-GMP is shown in stick representation. (B) Superposition of non-activated PleD chain A (blue) onto YdeH^{GGDEF} (orange). (C) Stereo close up view of the c-di-GMP coordination in the superposition of PleD (blue) onto YdeH^{GGDEF} (orange). The interacting residues are shown in sticks and the labeling corresponds to PleD.

Table 3.19: H-bonding between YdeH^{GGDEF} and c-di-GMP in the active site.

C-di-GMP	YdeH^{GGDEF}	Distance [Å]
N1	Asp182-$O_{\delta 1}$	2.7
N2	Asp182-$O_{\delta 1}$	3.1
N2	His178-O	3.1
N2	Asn173-$N_{\delta 2}$	2.8
N3	Asn173-$O_{\delta 1}$	3.2
O2'	Asn173-$O_{\delta 1}$	3.1
N7	Arg204-N_{H2}	3.4
O6	Arg204-N_{H2}	3.4

A structure of a GGDEF domain of *Xanthomonas campestris* (PDB: 3qyy) was solved very recently, which has c-di-GMP bound to the active site [151]. There, two GGDEF domains form a dimer by partial intercalation of the guanyl bases of two c-di-GMP molecules. In literature it is claimed that this c-di-GMP binding is an alternative inhibition mode to the domain crosslinking via the RxxD-motif, which is lacking in this enzyme. In comparison to YdeH^{GGDEF}, the binding of the guanyl base is the same in both structures, but the dimer arrangement is different.
In all three DGCs, the guanyl base of c-di-GMP binds in the identical hydrophobic pocket and the molecule is positioned through a H-bonding network consisting of the same residues. However, the dimer arrangement is different. The DGC structure from *X. campestris* shows a non-productive dimer, which is probably not of physiological relevance. The measured competitive K_I of 7 μM is too high in order to compete with GTP, which is present in millimolar concentrations in bacteria. However, in the structure of YdeH^{GGDEF} the dimer organization may resemble a physiological arrangement, because both active sites are linked with one c-di-GMP molecule in a symmetrical manner with the N-termini pointing in the same direction.

Substrate binding to the active site

In the full-length YdeH structure, the substrate analog GTPαS is bound to both active sites of the dimer in the same way. The guanyl base binds to the same hydrophobic pocket as the base of c-di-GMP in the YdeH^{GGDEF} structure. Furthermore, the base is positioned via H-bonding involving the same residues (Asn173, Asp182, Arg204 and His170). The ribose and the phosphate groups are interacting with Glu208, Arg281 and the backbone atoms of Lys170, Phe169, Ile166 and Arg168. Additionally, there are interactions with the residues Arg140 and Tyr205 of the other GGDEF domain. Furthermore, a magnesium ion, which is octahedrally coordinated by two oxygens of the β and γ phosphate of the GTPαS molecule and Asp165, Glu208 and Ile166, could be placed in the active site. The binding geometry is

Table 3.20: H-bonding between YdeH and GTPαS bound mainly to the A-chain in the active site.

GTPαS	YdeH	Distance [Å]	GTPαS	YdeH	Distance [Å]
N1	Asp182-$O_{\delta1}$	3.03	O1B	Lys170-N	3.3
N2	Asp182-$O_{\delta1}$	3.5	O1B	Phe169-N	3.3
N2	His178-O	3.1	O2B	Glu208-$O_{\eta2}$	2.9
N2	His178-$N_{\delta1}$	3.4	O2B	Ile166-O	2.7
N2	Asn173-$N_{\delta2}$	2.4	O2B	Arg168-N	2.6
N3	Asn173-$O_{\delta1}$	2.6	O2B	Phe169-N	2.5
N3	Arg140-$N_{\eta2}$ (chain B)	2.6	O3G	Asp165-$O_{\delta2}$	3.0
O6	Arg204-$N_{\eta2}$	3.0	O3G	Lys277-N_{ζ}	2.8
N7	Arg140-$N_{\eta1}$ (chain B)	3.3	O3G	Ile166-O	3.4
N9	Arg140-$N_{\eta2}$ (chain B)	3.2	O3G	Arg281-$N_{\eta1}$	3.4
O4'	Glu208-$O_{\eta1}$	3.3	O3G	Arg168-N	3.5
O2'	Arg140-$N_{\eta2}$ (chain B)	2.7	O1G	Arg168-N	3.0
O2'	Tyr205-OH (chain B)	2.9	O1G	Arg281-$N_{\eta1}$	2.5
O3A	Glu208-$O_{\eta1}$	2.9	O2G	Arg281-$N_{\eta2}$	3.4
O3A	Glu208-$O_{\eta2}$	3.1			

shown in Figure 3.36 and the distances of the interacting residues are listed in Table 3.20. In the full-length structure, of YdeH both active sites with bound GTPαS are facing each other. The distance between the α phosphate and the O3 of the GTPαS, which have to be linked during catalysis, is 10 Å (Figure 3.36). This means that the two GGDEF domains have to approach each other to allow the production of c-di-GMP. The zinc bound to the N-terminal CZB domain may cause the arrangement of the GGDEF domains to be in a non-productive conformation. The arrangement seen in YdeH full-length structure might represent a conformation, where the substrate has already bound, but needs minor rearrangement to initiate catalysis.

In the activated structure of PleD (PDB: 2v0n), the substrate analog GTPαS is also bound to the active site. The binding is very similar, only the conformation of the α phosphate is different. This results in a different position for Glu208 in YdeH. A superposition of the activated PleD structure onto YdeH is shown in Figure 3.36.

Inhibitory site

The I_p-site of DGCs is defined by the RxxD-motif, located at the opposite side of the active site on the loop between helix $\alpha2$ and β-strand $\beta2$. YdeH has a slightly different variant of this motif, RxxE. In the $YdeH^{GGDEF}$ structure dimeric c-di-GMP is bound to this I_p-site. Dimeric c-di-GMP has intercalated purine bases and forms two isologous intermolecular H-bonds between the guanine N1 and the phosphate groups, like it is known from small

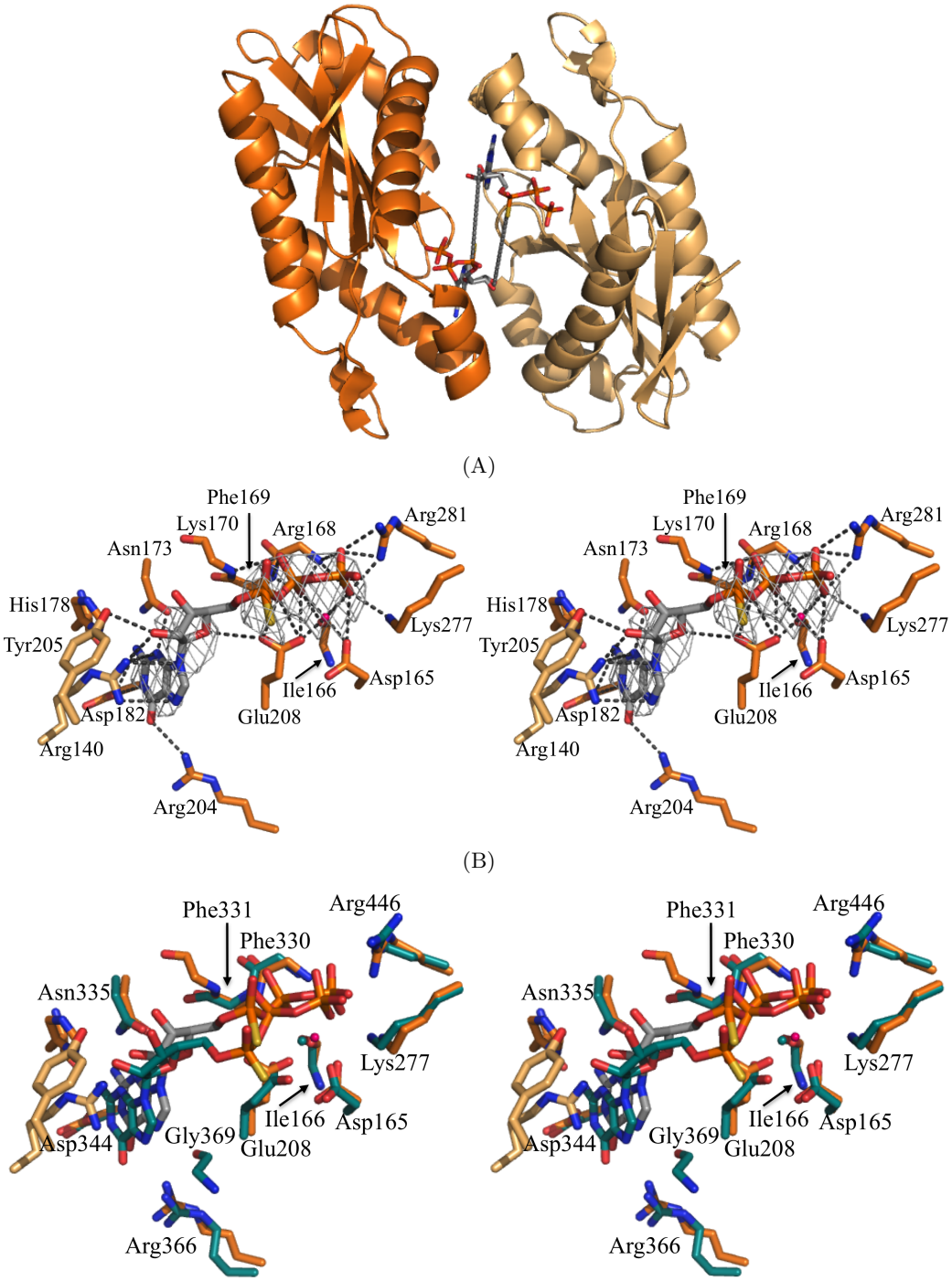


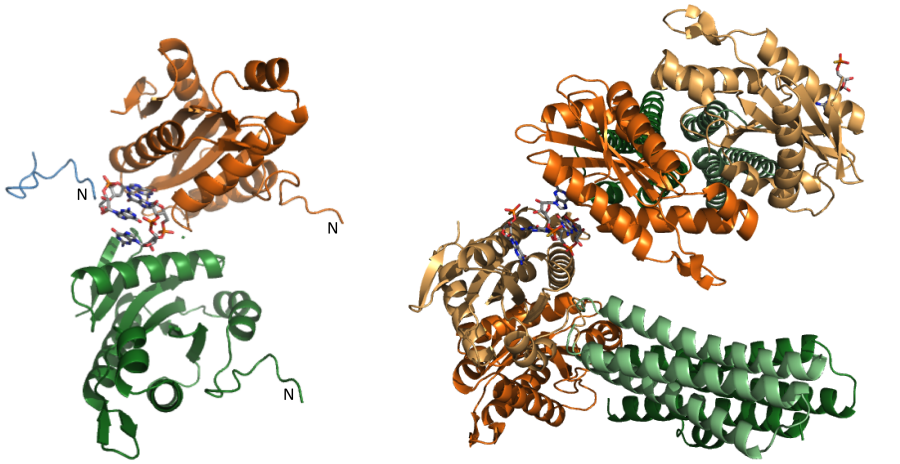
Figure 3.36: GTPαS binding to the active site in the full-length YdeH structure. (A) Ribbon representation of the two GGDEF domains. GTPαS is shown as sticks. The two active sites are facing each other. The two bonds, which have to be formed during catalysis, are marked as dashed lines. (B) Stereo view of the GTPαS coordination geometry at the active site. The interacting protein residues and GTPαS are shown in stick representation. The residues from chain A are colored in orange and the residues from chain B in light orange, respectively. The omit map for GTPαS is countered at 2.5 σ. (C) Stereo view of the GTPαS coordination in the superposition of activated PleD (green) onto YdeH^{GGDEF} (orange). The interacting residues are shown as sticks and are labeled for PleD.

molecule crystal structures of c-di-GMP [12, 13]. The c-di-GMP molecules are positioned with a strong H-bonding network involving Arg197 and Glu200 from the RxxE-motif and additional residues Arg224, Arg228, Asp198 and the backbone oxygen of Lys215 (Figure). The oxygen O1 from one phosphate group is involved additionally in the coordination of a magnesium ion together with Asp198 and four water molecules.
As it is known from other DGCs, binding of dimeric c-di-GMP to the I-site crosslinks two domains to prevent catalysis by domain immobilization [16]. In the YdeH^{GGDEF} structure, the c-di-GMP links two GGDEF domains, where the I_s-site is located on helix α4. Arg271 is interacting with both c-di-GMP molecules and the aromatic ring of Tyr274 stacks with one guanyl base. In addition, a third GGDEF molecule interacts weakly with c-di-GMP. The backbone nitrogen of Ala126 is in H-bonding distance to O11 of the phosphate group. The exact binding geometry is shown in Figure 3.37 and all distances of the interactions are listed in Table 3.21.
It is very likely that crystal packing is responsible for crosslinking of the third GGDEF molecule, which only participates with one H-bond of the backbone of Ala126, which is located on the N-terminus of the construct. The interaction with Arg271 and Tyr274 on helix α4 seems to be more relevant. The N-termini (Ile127) of the GGDEF domains are pointing in the same direction and are 31 Å apart. This might still occur in a dimer that is held together by a dimeric N-terminal domain and could therefore reflect an inhibition mode, which has to be proven by enzymatic experiments.

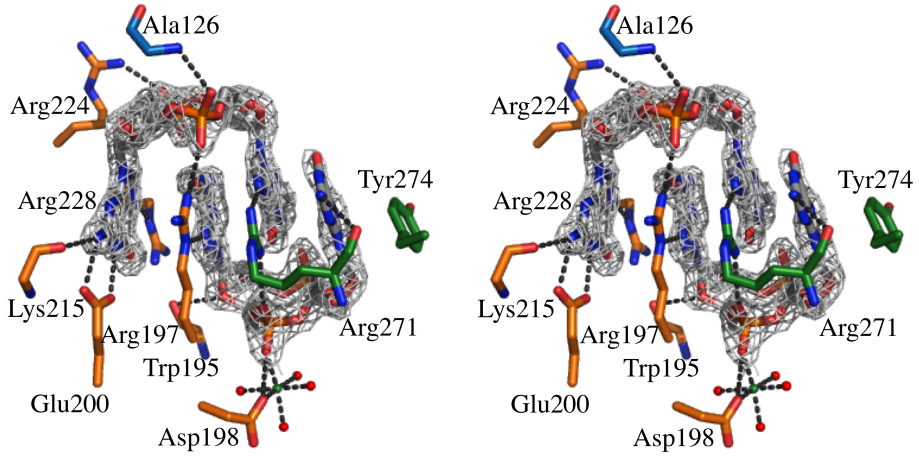
C-di-GMP was bound to the I-site in the full-length YdeH structure, as well. This was surprising, because neither c-di-GMP was added to the crystallization condition nor the protein used for crystallization contained c-di-GMP as confirmed by UV spectra. The bound c-di-GMP, which could be seen clearly in the initial electron density map, had been probably be slowly produced by YdeH from the substrate analog GTPαS during crystallization. This might explain, why crystals needed three days to occur and that they grew over a period of three weeks. Furthermore, the low amount of c-di-GMP could explain why only a c-di-GMP monomer was bound to the I-site and not a dimer as in all other known DGC structures. To compensate the limiting amount of c-di-GMP, additional GTPαS was bound to the I-site. This c-di-GMP/GTP arrangement does not use base intercalation, but the bases of c-di-GMP are stacked with GTPαS. Thereby, two GGDEF domains of symmetry related YdeH dimers are crosslinked. The I_p-site consists of residues Arg197 and Glu200 from the RxxE-motif and residues Arg228, Lys215 and Trp215, which position c-di-GMP and one GTPαS through a strong H-bonding network. The I_s-site on the symmetry related GGDEF domain is compromised of residues Tyr199, Glu200, Lys215, Arg224 and Arg228. These residues take part exclusively in the binding of the second GTPαS molecule. The domain crosslinking and the binding geometry of the I-site are shown in Figure 3.37 and all

interactions are listed in Table 3.21.
Since there are no interactions with c-di-GMP, the I_s-site is more likely a crystal artefact, which was due to the depletion of c-di-GMP in the crystallization solution. The GTPαS molecule bound to the I_p-site probably mimics one half of the missing c-di-GMP molecule. This would explain, why only the γ phosphate of the GTPαS has a defined position and is visible in the structure. On the other hand the α and β phosphate, which do not exist in c-di-GMP are not involved in the binding and are therefore not visible.

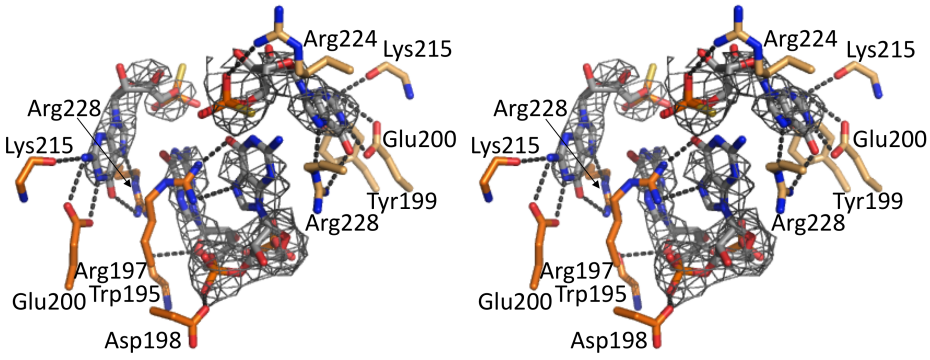
Comparing the binding of c-di-GMP to the I-site in the two YdeH structures, it becomes apparent that the c-di-GMP has a slightly different conformation (Figure 3.37). This happens, because only monomeric c-di-GMP is bound in the full-length structure of YdeH, which allows rotation of one guanyl base compared to the dimeric c-di-GMP in the YdeH^{GGDEF} structure. Arg197 is interacting with the intercalated base of the second c-di-GMP in the YdeH^{GGDEF} structure, while in the full-length structure Arg197 can adopt an other conformation and is involved in the binding of the guanyl base of monomeric c-di-GMP. The secondary I-sites are completely different, which is mainly caused by the binding of the additionally GTPαS molecules.


N
N
N

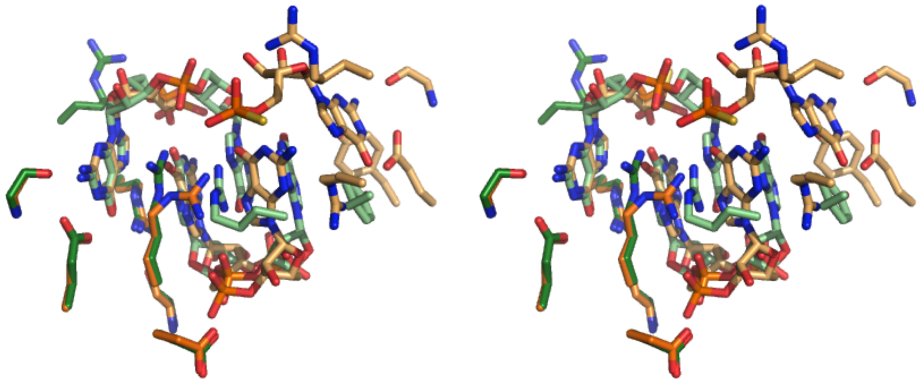
(A) (B)


Ala126
Arg224
Arg228
Tyr274
Lys215
Arg197
Arg271
Trp195
Glu200
Asp198
Ala126
Arg224
Arg228
Tyr274
Lys215
Arg197
Arg271
Trp195
Glu200
Asp198

(C)


Arg224
Lys215
Arg228
Lys215
Glu200
Tyr199
Arg228
Arg197
Glu200
Trp195
Asp198
Arg224
Lys215
Arg228
Lys215
Glu200
Tyr199
Arg228
Arg197
Glu200
Trp195
Asp198

(D)



(E)

Figure 3.37: The inhibitory site of YdeH. (A) Ribbon representation of c-di-GMP crosslinked GGDEF domains in the YdeH^{GGDEF} structure. The GGDEF domain with the c-di-GMP dimer bound to the I_p-site is shown in orange, the second GGDEF domain, where the c-di-GMP interacts with helix $\alpha 4$, is shown in green and the N-terminus of a third interacting GGDEF domain is shown in cyan. C-di-GMP is shown in sticks and a magnesium ion as green sphere. The N-termini of the orange and green GGDEF domains are pointing in the same direction and are 31 Å apart. (B) Ribbon representation of c-di-GMP crosslinked GGDEF domains in the YdeH full-length structure. The GGDEF domains are shown in orange and the CZB domains in green. Chain A and B are colored in dark and light colors, respectively. C-di-GMP is depicted as sticks. (C) Stereo view of the c-di-GMP coordination geometry at the I-site in the YdeH^{GGDEF} structure. The color code is the same as in (A). The interacting protein residues and c-di-GMP is shown in stick representation and the magnesium ion and the water molecules are represented as green and red spheres, respectively. The omit map for c-di-GMP is countered at $3\,\sigma$. (D) Stereo view of the c-di-GMP and GTPαS coordination geometry at the I-site in the full-length YdeH structure. The color code is the same as in (B). The interacting protein residues and c-di-GMP is shown in stick representation. The omit map for c-di-GMP is countered at $2.5\,\sigma$. (E) Stereo view of the comparison of the I-sites in the YdeH^{GGDEF} and the full-length YdeH structure. Interacting residues of YdeH^{GGDEF} and c-di-GMP are depicted in green and of full-length YdeH in orange.

In most of the other known DGC structures (PleD, PDB: 1w25, 2v0n; WspR, PDB: 3bre, 3i5c; and a DGC from *M. aquaeolei*, PDB: 3ign), c-di-GMP is bound to the I-site, as well [15–18]. In all these structures the primary I-site is very similar. It is defined by residues from the RxxD-motif and an additional arginine. Since YdeH has a glutamate in the I-site motif, which has one CH_2-group more than aspartate, the c-di-GMP dimer in YdeH is moved a little bit further from the protein (Figure 3.38). By that Arg228 can adapt an other conformation and is interacting with c-di-GMP only by the N_ϵ and not with the both N_η as in PleD for example. Nevertheless the exchange of the aspartate with a glutamate in the RxxD-motif does not impede c-di-GMP binding.

The secondary I-site compromises an arginine in the helix $\alpha 0$ like in the activated structure of PleD, WspR and in the DGC structure from *M. aquaeolei* or lies on the additional domains of the DGCs like in the non activated structure of PleD. All these binding modes have in common, that many arginines are involved, which can specifically interact with the O6-N7 edge of the guanyl bases. By crosslinking two domains, the two DGC domains are prevented from getting in close proximity to perform catalysis. This inhibition mode is therefore called 'inhibition by domain immobilization'.

The inhibition of YdeH can be induced by crosslinking the two DGC domains as it is seen in PleD, WspR and 3ign or by crosslinking the DGC domain to the CZB domain comparable to the non-activated structure of PleD. Additionally, crosslinking of two DGC domains might occur via the arrangement observed in the YdeH^{GGDEF} structure. To further investigate the inhibition of YdeH, activity assays had been performed (see section 3.1.10).

Table 3.21: H-bonding of nucleotides with YdeH^{GGDEF} and full-length YdeH in the inhibitory site.

H-bonding in YdeH^{GGDEF}			H-bonding in full-length YdeH		
C-di-GMP	YdeH	Distance [Å]	Nucleotide	YdeH	Distance [Å]
primary I-site					
503-O6	Arg197-$N_{\eta 2}$	2.9	c-di-GMP-O61	Arg197-$N_{\eta 2}$	2.5
503-N7	Arg197-N_{ϵ}	2.9	c-di-GMP-N7	Arg197-N_{ϵ}	2.9
503-O4'	Trp195-O	3.1	c-di-GMP-O4'	Trp195-O	3.4
503-O1P	Asp198-$O_{\delta 2}$	2.8	c-di-GMP-O1P	Asp198-$O_{\delta 2}$	2.7
505-O6	Arg228-$N_{\eta 2}$	2.7	GTP-O6	Arg228-$N_{\eta 2}$	2.9
505-N7	Arg228-N_{ϵ}	3.0	GTP-N7	Arg228-N_{ϵ}	3.1
505-N1	Glu200-$O_{\eta 1}$	2.6	GTP-N1	Glu200-$O_{\eta 1}$	3.2
505-N2	Glu200-$O_{\eta 2}$	2.7	GTP-N2	Glu200-$O_{\eta 2}$	3.0
505-N2	Lys215-O	2.8	GTP-N2	Lys215-O	2.7
505-O1P	Arg224-$N_{\eta 1}$	2.7			
505-O21	Arg197-$N_{\eta 2}$	2.9			
secondary I-site (x-y, x, z+1/3)			secondary I-site (-y, x-y+1, z+1/3)		
503-O2P	Arg271-$N_{\eta 1}$	2.8	GTP-O6	Arg228-$N_{\eta 2}$	3.3
503-N21	Arg271-$N_{\eta 2}$	3.0	GTP-N7	Arg228-N_{ϵ}	3.2
505-N7	Arg271-$N_{\eta 2}$	2.9	GTP-N1	Glu200-$O_{\eta 1}$	2.7
			GTP-N2	Glu200-$O_{\eta 2}$	3.1
			GTP-N2	Lys215-O	2.7
			GTP-N2	Tyr199-OH	3.2
			GTP-O3A	Arg224-$N_{\eta 1}$	3.3
tertiary I-site (x-y, x, z-2/3)					
505-O11	Ala126-N	2.8			

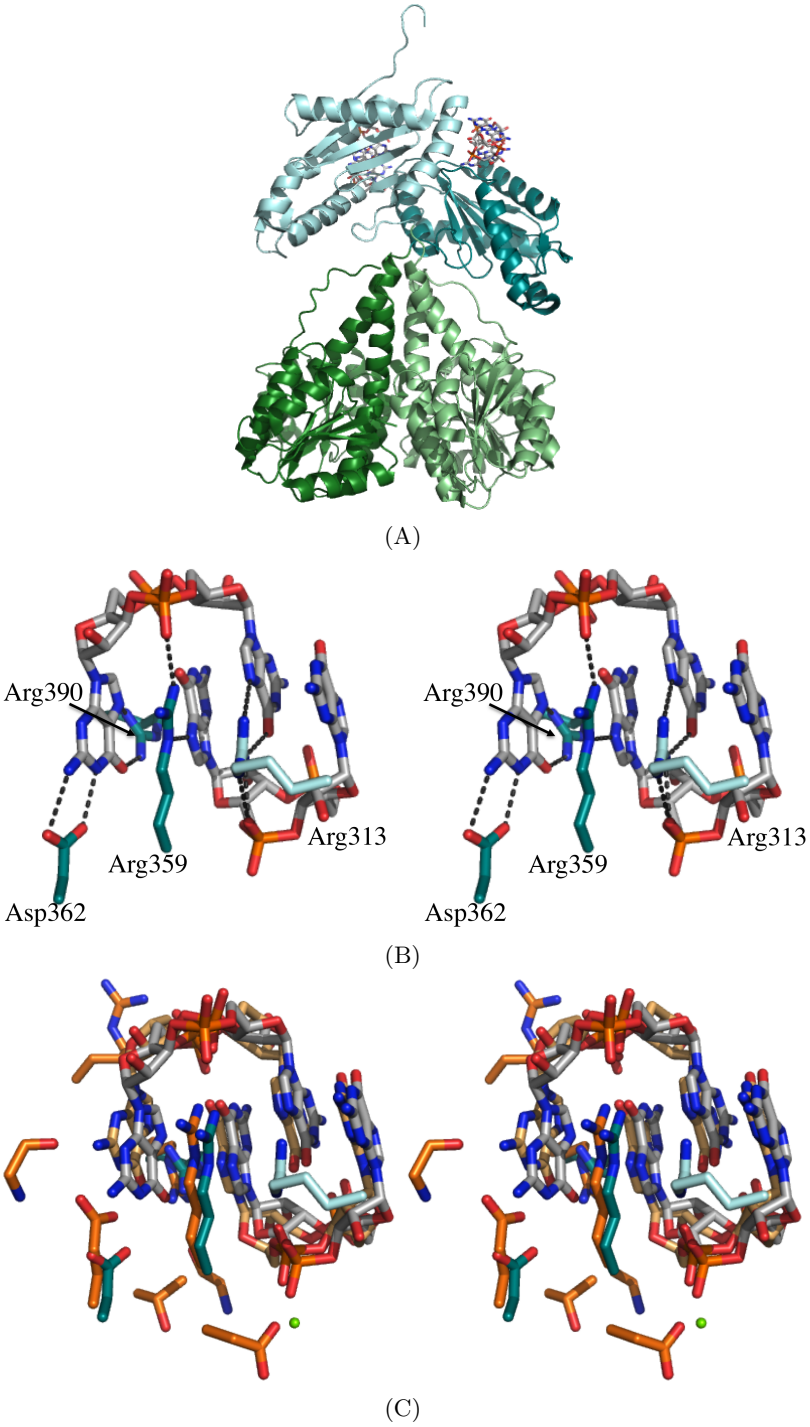


Figure 3.38: I-site of the activated structures of PleD (PDB: 2v0n) compared to YdeH^{GGDEF}. GGDEF domains of PleD are depicted in cyan and the N-terminal REC domains are shown in green. Chain A and B are colored in dark and light colors, respectively. (A) Ribbon representation of PleD, where c-di-GMP crosslinks the DGC domains.(B) Close up stereo view of the I-site of PleD. C-di-GMP and the interacting residues are shown as sticks. The coloring is the same as in (A). (C) Stereo view of the overlay of the I-site of PleD with the I-site of YdeH^{GGDEF}. In YdeH c-di-GMP is colored in light orange and the interacting residues of the I_p-site are depicted in orange and the residues of the I_s-site in light orange. The color code for PleD is the same as in (B).

3.1.8 Model of a catalytically competent YdeH-dimer

In all solved X-ray structures of DGCs the proteins are in a non-productive state meaning that the competent dimer is not known so far. In the full-length structure of YdeH, the substrate analog GTPαS is bound to both active sites. However, the dimer arrangement of the GGDEF domains is not in a competent state, because the intermolecular distance of 10 Å to form a bond between the 3'OH and the α phosphate of the GTPαS molecules is far too long (Figure 3.36).
In the structure of YdeH^{GGDEF}, the two active sites are crosslinked via c-di-GMP (Figure 3.34). GTPαS was modeled in the active sites of this YdeH^{GGDEF} dimer by superimposing with the YdeH full-length structure and revealed, that this dimer arrangement can not be a productive state as well (Figure 3.39). The distance of 9.3 Å between the 3'OH and the α phosphate of the GTPαS molecules is still too long to form a bond and is not in-line with the leaving group. The base of a GTPαS molecule overlays very well with the bases of c-di-GMP, but the position and conformation of the ribose and the α phosphate moiety is completely different. Thus, the position of the base is not changed during catalysis and always stays in the same binding pocket. The ribose and the α phosphate adopt different conformations during catalysis to end up in the relatively rigid conformation of the c-di-GMP macrocycle. To allow this conformational change, the domains of the GGDEF dimer have to rearrange too in order to enable catalysis.
A model was generated to mimic the competent state of YdeH, setting the distance of the forming bond between the 3'OH and the α phosphate of the GTPαS molecules to 3 Å and the attack occurs in-line (Figure 3.39). In this model no clashes appeared, but the N-termini (Asn130) of both GGDEF domains are 14.2 Å apart from each other. This distance is probably too far to allow dimerization via the N-terminal CZB domains. However, the generated model can serve as a starting point for further model refinement, because already smallest changes of the torsion angles of the glycosidic bond and the bond between C5'-O5'-Pα of the GTP cause a big impact on the dimer arrangement of the DGC domains. By adjusting these torsion angles, the model can be refined in such way, that the N-termini of the GGDEF domains can come closer together in order to allow the dimerization via the CZB domains.
Independent of how the exact competent state of YdeH looks like, a reorientation of the GGDEF domains relative to each other has to occur to switch from the substrate bound state to the product bound state. To accomplish this transition, flexibility is required between the GGDEF and the CZB domain.

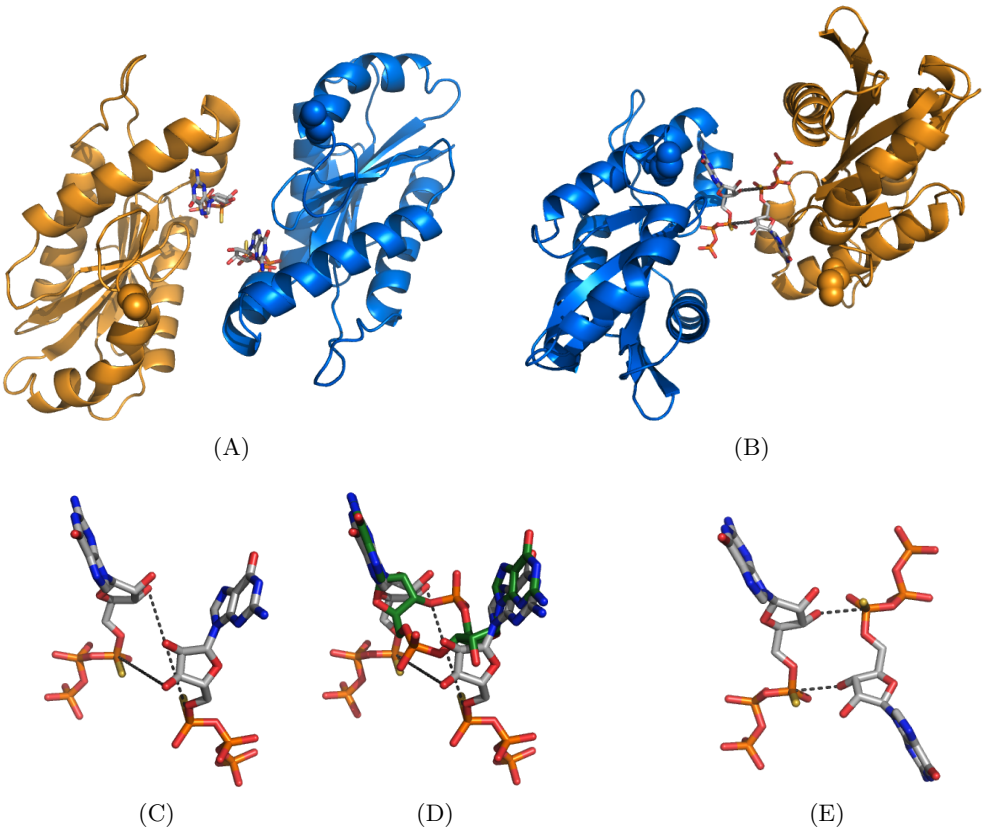


Figure 3.39: (A) Dimer arrangement seen in the YdeH^{GGDEF} structure applied on the DGC domains as found in the full-length YdeH structure. One monomer is shown in orange and one in blue. The residue Asn130 representing the C-terminus is shown as sphere and GTPαS is depicted as sticks. (C) Close up view of (A). The forming bonds are marked with dashed lines. (D) Close up view of (A) overlayed with c-di-GMP (green) as it is bound in the YdeH^{GGDEF} structure. (B) Model of the competent dimer of YdeH. The forming bond has a distance of 3 Å and the O3' is in-line with the leaving group. One monomer is shown in orange and one in blue. The residue Asn130 representing the C-terminus is shown as sphere and GTPαS is depicted as sticks. (E) Close up view of (B). The forming bonds are marked with dashed lines.

3.1.9 Summary of the YdeH structures

Three structures of YdeH could be determined including the full-length structure at a resolution of 3.9 Å and the structures of the CZB domain and the GGDEF domain of YdeH at 2.2 Å and 1.8 Å resolution, respectively.
In the structure of full-length YdeH, substrate binding to the active site could be identified, which is similar to the binding in the activated structure of PleD. However the dimer arrangement of the two GGDEF domains is different. In YdeH the two active sites are facing each other, but are still not in a competent conformation, because the substrate molecules are too far apart from each other to perform catalysis.
Product binding to the active site could be shown in the structure of $YdeH^{GGDEF}$. Here, c-di-GMP crosslinks the two active sites. The guanyl bases of c-di-GMP bind to the same pocket as the substrate, but the ribose and the α phosphate moiety adapt different conformations and are not involved in protein interactions. In this structure, the dimer arrangement of the two GGDEF domains is stabilized via two salt bridges, consisting of the conserved residues Arg140 and Asp174. To confirm the importance of this GGDEF arrangement enzymatic experiments will have to be performed.
In the presented structures of full-length YdeH and $YdeH^{GGDEF}$, c-di-GMP was bound to the I-site of YdeH. The dimeric binding of c-di-GMP to the I_p-site in the $YdeH^{GGDEF}$ structure was very similar to other known structures of DGCs, where c-di-GMP is bound to the I-site, as well. Since YdeH has a variation in the I_p-site motif (RxxD), where the aspartate is exchanged to a glutamate, it could be shown that this does not impede the c-di-GMP binding. The I_s-site in the $YdeH^{GGDEF}$ structure is completely different from the other known DGC structures. The seen I_s-site is located on helix α4 and consists of Arg271 and Tyr274. Since the N-termini of this crosslinked DGC domains are pointing in the same direction, this I_s-site has to be analyzed in enzymatic experiments. Furthermore a I_s-site on the N-terminal CZB domain has to be considered.

The N-terminal CZB domain of YdeH shows a new fold consisting of a four helical bundle. One interesting finding was that the CZB domain of YdeH harbors a zinc-binding site compromised of three histidines and one cysteine, which mediates contact between the helices and fix them in a rather rigid arrangement. Consequently the question arises, if zinc is the input signal for YdeH and can therefore regulate the diguanylate cyclase activity of YdeH upon binding to the zinc binding site in the N-terminal domain. This was tried to answer by enzymatic investigations shown in the next section.

3.1.10 Enzymatic activity and regulation of YdeH

YdeH shows DGC activity

To test if YdeH shows diguanylate cyclase activity, 2 μM YdeH was incubated with 100 μM GTP in 50 mM Tris-HCl, pH 8.0, 50 mM NaCl and 5 mM $MgCl_2$. The reaction was stopped after 2 and 15 min by boiling 2 min and analyzed by ion-exchange chromatography as described in section 2.3.3. The elution of the reaction was compared to GTP and c-di-GMP elution profiles. As shown in Figure 3.40, the peak of the substrate GTP decreased, while the c-di-GMP peak increased during the reaction. The reaction product of YdeH overlays with the c-di-GMP-standard peak and mass analysis confirmed the product to be c-di-GMP. Thus, it could be shown that YdeH indeed acts as a diguanylate cyclase *in vitro*.

To determine the kinetic parameters of YdeH, enzymatic activity of YdeH was measured with the Baykov assay as described in section 2.3.3. For this, 2 μM zinc-free YdeH was incubated with different concentrations of GTP (5-100 μM) and the initial velocities were measured. Maximal velocity of 5.5 min^{-1} was measured up to a GTP concentration of 30 μM (Figure 3.41). Below 20 μM GTP, no linear initial velocity could be measured, even if the enzyme concentration was decreased and the reaction time was shortened. With a enzyme concentration, which would result in a initial linear velocity such small amounts of product would be produced, which is under the detection limit of the assay. Only an upper limit for the K_M of 5.8 μM could be determined. v_{max} was calculated to 5.5 min^{-1}.
These values for K_m and v_{max} are in the same range as already described for other DGCs [16, 31, 103]. The velocity of YdeH is very slow for an enzyme. This might be due to not optimal reaction conditions or to the absence of an additional factor that is required to stimulate activity. On the other side, catalysis is a quite complicated process, where two domains have to approach each other in the right orientation to form the complete active site. The requirement for these specific conformation could also limit the velocity of the catalyzed reaction.

The CZB domain is necessary for YdeH activity

To investigate the importance of the N-terminal CZB domain of YdeH, activity measurements of full-length YdeH were compared with $YdeH^{GGDEF}$, in which the complete CZB domain is missing. Enzymatic activity was measured with the Baykov assay with 2 μM enzyme and 100 μM GTP in 50 mM Tris-HCl, pH 8.0 , 50 mM NaCl and 5 mM $MgCl_2$. At time points of 1, 2, 3 and 4 min the production of pyrophosphate was measured. Under these condition, no activity of $YdeH^{GGDEF}$ was measurable. This indicates, that the GGDEF domain alone is much slower or even not active compared to full length YdeH.

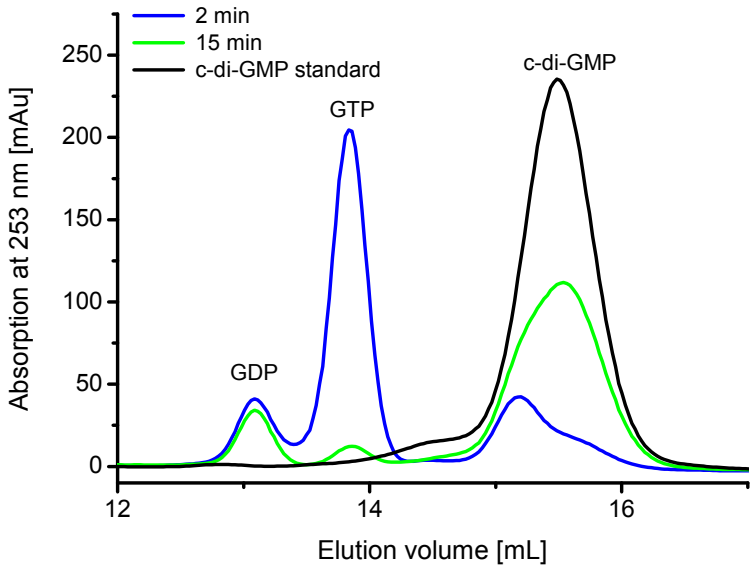


Figure 3.40: Time course of the reaction catalyzed by YdeH. The samples were analyzed by ion-exchange chromatography. In black, a run of a c-di-GMP standard is shown. The left most small and constant peak corresponds to a GDP contamination in the commercial GTP batch.

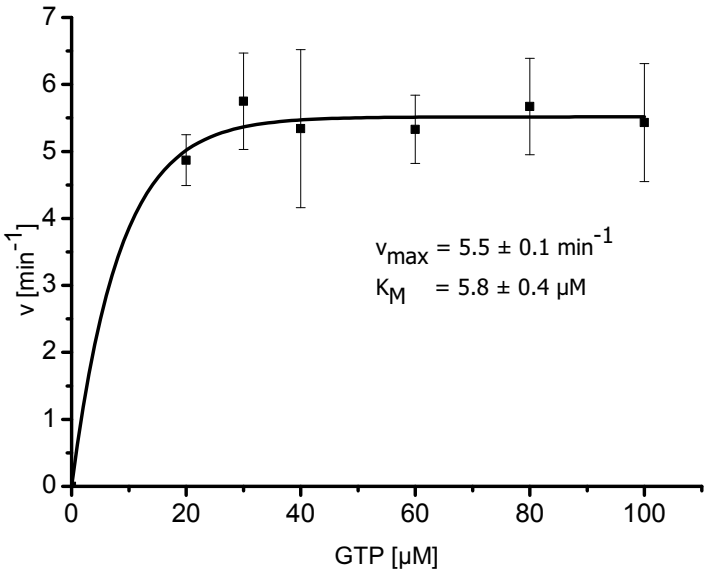


Figure 3.41: Initial velocity at different GTP concentrations. K_M and v_{max} values were determined by fitting the Michaelis-Menten curve. Error bars are standard deviations.

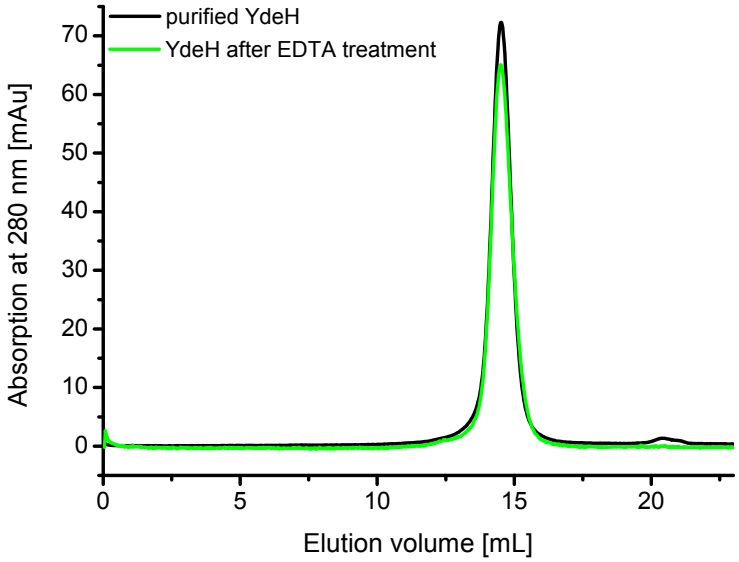


Figure 3.42: Gel filtration runs with YdeH samples before and after treatment with EDTA. The column S200/10/30 with 20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 50 mM arginine and 50 mM glutamate as running buffer was used.

Furthermore, reaction setups with 20 and 100 μM YdeH^{GGDEF} and 500 μM GTP with time points at 6, 23 and 26 h were performed. Under these conditions very little activity could be measured resulting in a turnover of $1.8 \times 10^{-5}\,min^{-1}$ and $2.9 \times 10^{-5}\,min^{-1}$, which is 250000 times slower compared to full-length YdeH. This shows, that the DGC domain alone is not able to perform catalysis in a reasonable timescale. Thus, the dimer formation of the N-terminal CZB domain is essential to perform the catalytic reaction, as it raises the likelihood of the productive encounter of two GGDEF domains.

YdeH is inhibited by zinc

To investigate the effect of zinc on YdeH activity, enzyme that is completely devoid of zinc has to be produced for comparison. Therefore, YdeH was incubated with EDTA over night as described in section 2.3.1. To control, if YdeH is still well folded after zinc removal and to remove the EDTA, gel filtration was performed. As it is shown in Figure 3.42, the protein behaves the same as before EDTA treatment, meaning that YdeH is stable and well folded without zinc. This was a first indication, that zinc has no important structural role.
To prove that YdeH is zinc-free after EDTA treatment, a PAR assay was performed. With this assay it could be demonstrated, that all zinc was released from the protein. For comparison, a protein sample before EDTA treatment was analyzed resulting in a zinc content

of 40 %. This indicates, that YdeH has a high affinity binding site for zinc, because zinc is still bound after all the purification steps, in which no zinc was added. The removal of zinc with EDTA was possible because a high excess of EDTA and a long incubation time was used.

To elucidate the effect of zinc on YdeH activity, 50 μM zinc free YdeH was incubated with different zinc concentrations (0-150 μM). To measure the amount of zinc, which has bound to YdeH, part of the sample was analyzed with the PAR assay. A small variation of the PAR assay was used, in which the mercury derivative was omitted, in order to measure the remaining zinc in solution. To compare zinc binding with activity of YdeH, a second part of the sample was used to perform a Baykov assay. Therefore 2 μM zinc treated YdeH sample was incubated with 100 μM GTP in 50 mM Tris-HCl, pH 8.0 , 50 mM NaCl and 5 mM $MgCl_2$. In Figure 3.43, the activity and the amount of zinc bound to YdeH was plotted against the total zinc concentration added. At the beginning a linear increase of zinc binding to YdeH takes place until the protein concentration of 50 μM is reached, meaning that up to 50 μM all added zinc is completely bound to YdeH. Above 50 μM further zinc is bound by YdeH, which probably results from unspecific binding to the protein surface. The divalent cation zinc can function as a crosslinker, which leads to protein aggregation. This could be shown in gel filtration experiments (Figure 3.44), where YdeH, incubated with an excess of zinc, forms higher aggregates.
An increase of zinc bound to YdeH results in a reduced enzymatic activity. Hardly any detectable activity is measurable at equimolar zinc and protein concentration indicating that YdeH is inhibited by zinc.

This experiment showed, that zinc inhibits the activity of YdeH. The zinc binding is very strong with a K_I below 50 μM, because until the YdeH concentration of 50 μM is reached nearly every zinc added is bound by the protein. In order to determine the exact K_I for the zinc binding to YdeH at least ten times lower protein concentration than the K_I have to be used. Since this is not possible with the employed Baykov assay, competition experiments were performed. For this competition assay, where YdeH and the chelator EDTA compete for the zinc, 0.2 μM YdeH was incubated with 0.2 μM $ZnCl_2$ and different EDTA concentrations (0.1 nM-1 mM). The activity was measured with the Baykov assay with 100 μM GTP in 50 mM Tris-HCl, pH 8.0, 50 mM NaCl and 5 mM $MgCl_2$. The activity was plotted against the EDTA concentration, as it is shown in Figure 3.45. The data was fitted with an exact mathematical model [107] describing competitive binding of EDTA and YdeH to zinc. The K_D (EDTA/Zn) was set to 1×10^{-14} M to achieve a stable fitting. The determined values were than adjusted to the real K_D (EDTA/Zn) of 4.79×10^{-16} M [152], which resulted in a K_I (YdeH/Zn) of 3.4×10^{-17} M.

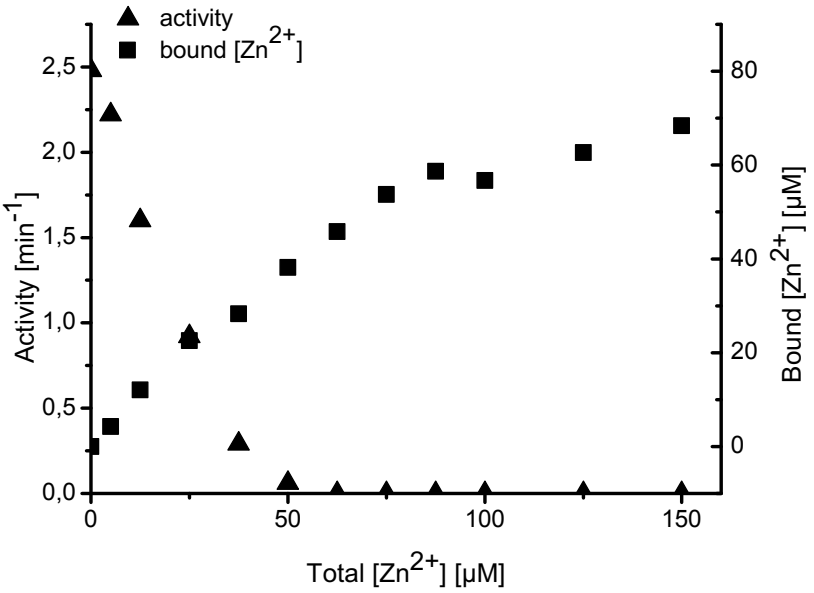


Figure 3.43: Inhibition of YdeH by zinc. The total zinc concentration is plotted against the amount of zinc, which is bound to YdeH (■), and against the activity of YdeH (▲). The amount of free zinc was measured with the PAR assay and the activity was determined with the Baykov assay.

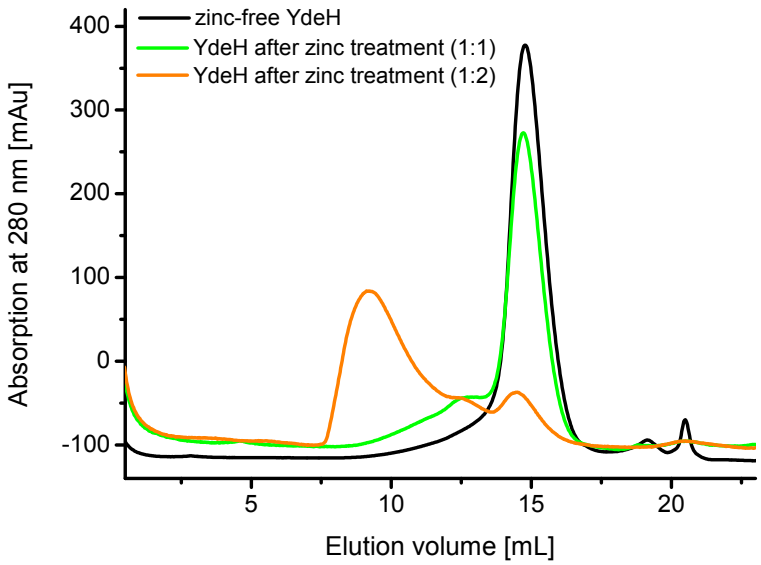


Figure 3.44: Gel filtration runs of YdeH incubated with different amounts of zinc.The column S200/10/30 with 20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 50 mM arginine and 50 mM glutamate as running buffer was used.

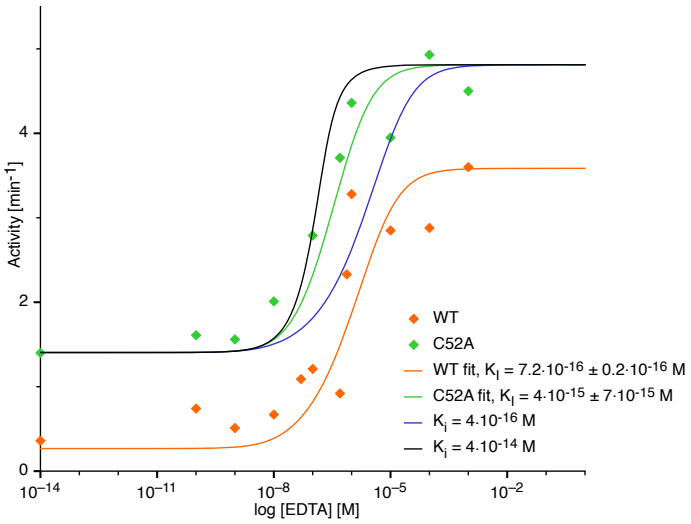


Figure 3.45: YdeH DGC activity in the presence of a zinc binding competitor (EDTA). Inhibition of wildtype YdeH activity (orange) and YdeH-C52A activity (green) is relived by zinc sequestration to EDTA. Protein and zinc concentrations were 0.2 μM. The data were fitted with an exact mathematical model [107] describing competitive binding of EDTA and YdeH to zinc. The K_D (EDTA/Zn) was set to 1×10^{-14} M to achieve a stable fitting yielding in K_I (YdeH/Zn) of 7.2×10^{-16} M for wildtype YdeH and in a K_I (YdeH-C52A/Zn) of 4×10^{-15} M for the C52A mutant. For the C52A mutant an upper and lower limit with a K_I of 4×10^{-14} M and 4×10^{-16} M was additionally plotted in the graph.

Noteworthy the activity is not completely abolished at small EDTA concentrations. This can either result from incorrect determined YdeH or $ZnCl_2$ concentrations, or the affinity of YdeH for zinc is not high enough to bind an equimolar amount of zinc. Furthermore, the residual activity may be due to the encounter of GGDEF domains from different dimers.

Based on this exquisitely small K_I value, it appears very unlikely that the inhibition of YdeH is due to nonspecific zinc binding. Nevertheless to corroborate the idea, that zinc binding to the CZB domain of YdeH controls its enzymatic activity, it was tried to generate zinc binding mutants of YdeH. It was attempted to individually change all four zinc coordinating residues His22, Cys52, His79 and His83. However, despite several repeated attempts using two different cloning strategies, it was not possible to generate mutations in pET28a/YdeH that would exchange any of the zinc coordinating histidines to leucines. The only variant that could be generated was the C52A mutant.
To analyze the zinc content after purification of YdeH-C52A, a PAR assay was performed. Surprisingly, in contrast to the wildtype there was no zinc detectable, which already indicates a weaker affinity of this mutant for zinc.
To investigate the influence of zinc binding on the activity of YdeH, the same competition

experiment was performed with the C52A mutant. As shown in Figure 3.45, the residual activity of the mutant at low EDTA concentrations is much higher than for the wildtype. In addition, the inflection point of the fitted data curve is shifted to lower EDTA concentration compared to wildtype. This results in a ten times higher K_I of 2×10^{-16} M. This indicates clearly, that Cys52 is involved in zinc-binding as it is seen in the crystal structures. To demonstrate that the fitted K_I is in the right range, an upper and lower limit with a K_I of 2×10^{-15} M and 2×10^{-17} M was additionally plotted in the graph.

The measured inhibition of YdeH by zinc is in the femtomolar range and is comparable to affinity constants of the zinc sensor proteins ZntR and Zur from *E. coli* that are zinc dependent transcriptional regulators [54, 64]. This remarkably high affinity for zinc is required to allow full occupation of zinc containing proteins, because the free zinc concentration in the bacterial cell is very limited due to the reactivity of zinc as a potent Lewis acid. This very low concentration of free zinc and the high affinity binding sites in zinc containing proteins, however, have important consequences for the kinetics of zinc binding to proteins. Ligand exchange mechanisms for zinc transfer from occupied to unoccupied bindings sites are necessary to circumvent the kinetic limitations of dissociation/assosiation mechanisms, which are caused by the virtual absence of free zinc ions. A mechanism of zinc transfer between alternative binding partners via ternary complexes was proposed, where a considerable fraction of individual ligands of a zinc binding site might be in a rapid equilibrium of bound and unbound state. This results in a transient availability of binding sites for competing acceptor molecules. The transfer via zinc-bridged complexes can be many orders of magnitude faster and does not depend on free zinc ions for rapid mediation of zinc signals [153]. The unstructured loop region close to the zinc binding cysteine in the N-terminal domain of YdeH is ideally suited for this fast zinc exchange. This findings strongly supports the idea that zinc binding controls YdeH activity, rather than playing a structural role.

YdeH is a zinc sensor and not regulated by a redox mechanism

As the activity of YdeH is regulated by zinc, the question arises if YdeH is a zinc sensor and therefore acts in response to the zinc level. Another possibility might be that YdeH is regulated by a redox mechanism, that allows the control of YdeH activity via the zinc content as a consequence of the redox state of the coordinating cysteine residue. The sulfur donor atoms of cysteine ligands can be oxidized and reduced again with concomitant dissociation and association of zinc. *In vivo* oxidative stress can be caused by macrophages, that would lead to activated YdeH. The resulting increased c-di-GMP level would induce biofilm formation.

To test this possibility, the affinity of YdeH for zinc should be decreased after treatment with an oxidizing agent. This would result in zinc release and activation of the enzyme. The following different oxidizing agents were tested in an excess of 5, 50, 500 and 5000: iodacetamide, N-ethylmaleimide, H_2O_2, KSCN, $NaAsO_2$, 4-chloromercuribenzoic acid, oxidized glutathione, Paraquat, pyocyanin, pyocianin/NADH, phenylselenyl chloride and glutathione reductase/oxidized glutathione. Some of these different reagents are used in laboratory praxis to oxidize cysteines and others are naturally occurring oxidizing agents. The incubation was done with 1 μM of YdeH and 1 μM $ZnCl_2$ in 50 mM Tris-HCl, pH 8.0, 50 mM NaCl and 5 mM $MgCl_2$. Afterwards the samples were tested for DGC activity with the Baykov assay by adding 100 μM GTP. The result for all different oxidizing reagents was the same: the protein has completely lost its enzymatic activity. This was most likely a result of unspecific oxidation, which leads to protein aggregation.
Furthermore the same experiments were performed with the C52A mutant of YdeH, which also resulted in a entire loss of DGC activity. This supports the idea of protein aggregation as a result of unspecific oxidation.

Since the loss of activity also occurred under rather mild conditions by using *in vivo* occurring systems like glutathione reductase/oxidized glutathione, it is unlikely that the activity of YdeH is regulated via a redox mechanism. It seems to be more reasonable, that YdeH is directly regulated by zinc. This finding is also supported by *in vivo* data. It could be shown, that oxidative stress has no effect on cellular c-di-GMP concentration and on the PgaD level, which was shown to be controlled by c-di-GMP produced by YdeH [47]. Furthermore, external zinc represses PgaD levels via YdeH. Mutations in the zinc-binding site of YdeH results in constitutive high level of PgaD, which is not responding to external zinc. In a ΔYdeH strain, the PgaD level is constantly low and is again not susceptible to external zinc (A. Böhm, unpublished data).

The salt bridge Arg140-Asp174 is important for YdeH activity

To confirm the dimer arrangement observed in the $YdeH^{GGDEF}$ structure with bound c-di-GMP (Figure 3.34), the effect of the salt bridge Arg140-Asp174 on DGC activity of YdeH was investigated. Therefore, two mutants were generated, whereby first the aspartate 174 was exchanged to an arginine and in a second mutant the arginine 140 was additionally mutated to a aspartate.
Purified and zinc-free YdeH-D174R was analyzed with the Baykov assay for DGC activity. In the beginning, the same standard conditions as for wildtype YdeH were used: 2 μM YdeH-D174R and 100 μM GTP in 50 mM Tris-HCl, pH 8.0, 50 mM NaCl and 5 mM $MgCl_2$. No activity could be detected under this conditions. Then, the enzyme concentration was

increased to 10 μM and the reaction time was increased from 4 to 120 min. Under these conditions, a turnover of $4 \times 10^{-4}\,\mathrm{min}^{-1}$ was measured. The velocity of this mutant is 14000 times slower than for the wildtype. The exchange of a negative charged amino acid to a positive one prevents the formation of the salt bridge and induces an electrostatic repulsion. Thus, the formation of the salt bridge Arg140-Asp174 may be important for YdeH to perform the catalysis. This finding confirms the dimer arrangement as it is seen in the $\mathrm{YdeH}^{\mathrm{GGDEF}}$ structure.
In a next step, the salt bridge should be restored with the double mutant YdeH-R140D-D174R. Because the GGDEF domains form a symmetric dimer, the formation of the salt bridge should be possible. Unfortunately, no soluble protein could be expressed from this second mutant YdeH-R140D-D174R. This is astonishing, because the overall charge of the protein was not changed and no big conformational changes were expected.

YdeH shows product inhibition

Several DGCs show inhibition by the reaction product c-di-GMP [15, 17, 154]. The first crystal structures of DGCs led to the hypothesis, that inhibition is due to the crosslinking of the GGDEF domain to another domain, so that the two GGDEF domains can not approach each other to perform catalysis. This inhibition by domain immobilization model was later confirmed by enzymatic experiments [16, 31]. Another proof for this mechanism came from two structures of WspR, where one has a c-di-GMP dimer bound to the I-site and the second is nucleotide free. By comparison it could be shown, that the structures of the GGDEF domains are identical in both states. The active site loop is very similar and independent from the c-di-GMP binding to the GGDEF domain and therefore a communication between the $\mathrm{I_p}$-site and the active site could be excluded to cause the noncompetitive product inhibition [18].
To investigate the mechanism of product inhibition, YdeH was analyzed by enzymatic experiments. For this, 1 μM zinc-free YdeH was incubated with different c-di-GMP concentrations (0-2 mM) in 50 mM Tris-HCl, pH 8.0, 50 mM NaCl and 5 mM $\mathrm{MgCl_2}$ and was analyzed with the Baykov assay by adding 100 μM GTP. As it is shown in Figure 3.46, YdeH is inhibited by c-di-GMP with a K_I of 58 μM. To exclude that the measured inhibition is not a competitive inhibition, the same experiment was performed with a different GTP concentration of 500 μM GTP. This resulted in the same K_I value, which proves that YdeH is non-competitively inhibited by c-di-GMP. The measured inhibition constant of 58 μM is far above the physiological concentration of c-di-GMP *in vivo*, which means that the measured inhibition is probably not physiological. One explanation for this high K_I could be the difference in the RxxD-motif compared to other studied DGCs. The exchange from an aspartate to a glutamate in YdeH might result in higher K_I, or the difference in K_I is located in the nature of the secondary so far not reliably identified I-site.

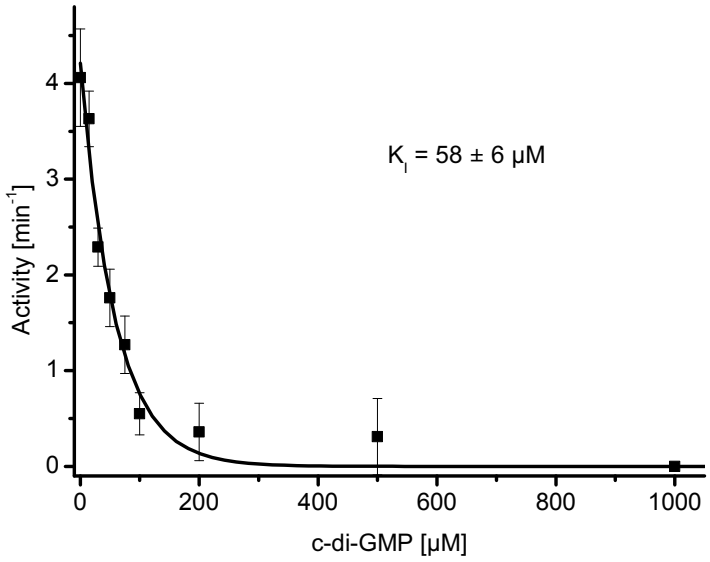


Figure 3.46: Product inhibition of YdeH. The activity is plotted against c-di-GMP concentration. The K_I was determined to be 58 μM. Error bars are standard deviations.

As in the YdeH^{GGDEF} structure, a c-di-GMP mediated crosslinking between two GGDEF domains is seen (Figure3.37). It was tested, if this domain interaction is involved in inhibition of YdeH. Therefore, the interacting residues of the possible I_s-site Arg271 and Tyr274 were mutated to alanines. Surprisingly, the double mutant showed no enzymatic activity. As the protein was highly expressed, very soluble and showed no aggregation on gel filtration experiments, the loss of activity is rather unlikely due to misfolded protein. One explanation might be, that the mutated residue Tyr274 is important to form the competent dimer to allow catalysis. As it is shown in Figure 3.47, this residue is nearly in H-bonding distance with its symmetry mate in the dimer interface of the two GGDEF domains in the YdeH^{GGDEF} structure. During catalysis, in which the two DGC monomers have to be closer together, the Tyr274 can be important for forming a tighter dimer. In addition, both mutated residues Arg271 and Tyr274 are located at the dimer interface of the two GGDEF domains in the full-length structure of YdeH. Arg271 is 9.0 Å away from Asp174 and Tyr274 is about 9.0 Å away from Lys170, which is not completely visible in the structure. These two pairs might also form salt bridges in the competent dimer during catalysis, when the two GGDEF domains are in close proximity. To further investigate the role of Arg271 and Tyr274, the single mutants would have to be generated and analyzed for activity and product inhibition.

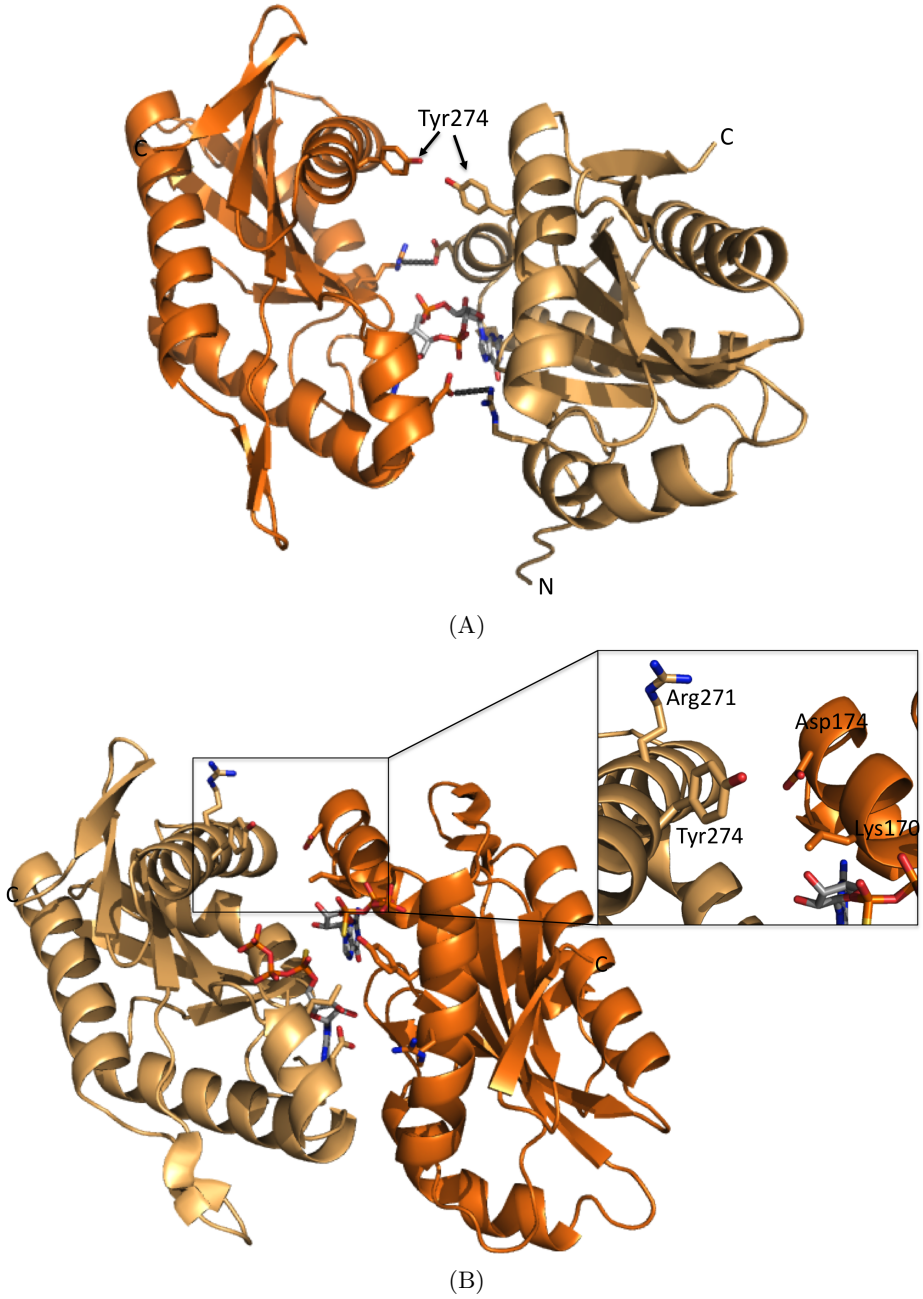
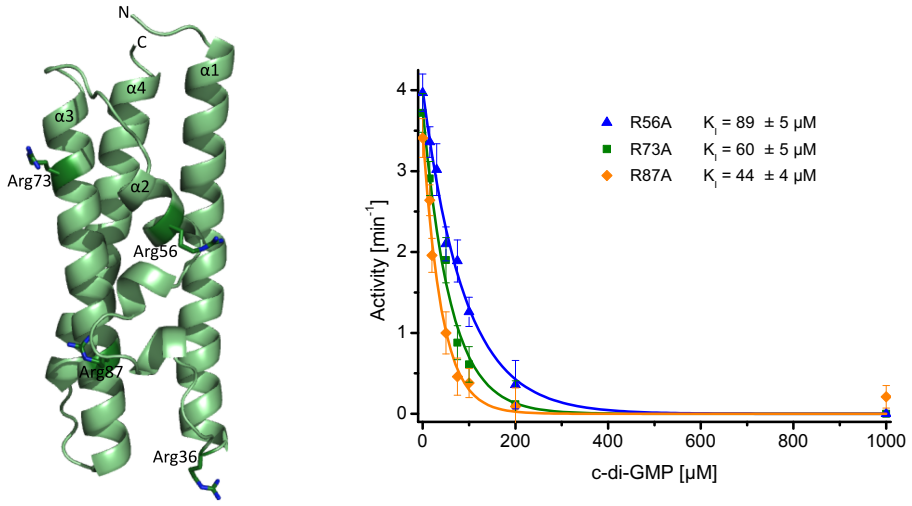


Figure 3.47: Role of Tyr274 and Arg271 in the dimer interface of YdeH. (A) Ribbon representation of YdeH^{GGDEF}. The two monomers are colored in light and dark orange, respectively. C-di-GMP and the interacting residues in the dimer interface (Arg140 and Asp174) are shown as sticks as well as the Tyr274, that are 4.0 Å apart and therefore close to H-bonding distance. (B) Ribbon representation of the GGDEF domains of full-length YdeH. The two monomers are colored in light and dark orange, respectively. C-di-GMP and the residues in the dimer interface that could interact with Arg271 and Tyr274, if the dimers move closer together are shown as sticks. In the right corner, a zoomed view of the possible interacting residues is depicted.



(A) (B)

Figure 3.48: Investigation of a secondary I-site on the CZB domain of YdeH. (A) Ribbon representation of the CZB domain of YdeH. All arginines on the surface are represented as sticks and labeled. (B) Product inhibition of the R56A, R73A and R87A mutants of YdeH. To determine the K_I the activity was plotted against c-di-GMP concentration and fitted with a exponential decay function. Error bars are standard deviations.

Furthermore, the presence of a I_s-site located on the CZB domain of YdeH, comparable to the N-terminal REC domain of PleD [15] was investigated. Since arginines are always involved in c-di-GMP binding, the arginines on the surface of the CZB domain were chosen for a possible interaction with the I-site of the GGDEF domain (Figure 3.48). Of the four arginines of the CZB domain, Arg36 can be excluded, because this residue is to far away to interact with the I_p-site on the GGDEF domain. For the other three arginines, Arg56, Arg73 and Arg87, single mutations were generated by exchanging the arginines to alanines. All three mutants were tested for product inhibition by incubating the zinc-free protein with different concentrations of c-di-GMP (0-1 mM) and checking for activity with the Baykov assay, as it was done for the wild-type. For all three mutants, the calculated K_I values lay in the same range as for wild-type YdeH (Figure 3.48). This supports that the tested arginines are not involved in a I_s-site. It cannot fully excluded that a I_s-site is present in the CZB domain comprised of other residues, however, this is rather unlikely as arginines are involved in all so far known binding modes of c-di-GMP.

With the mutants tested so far, no I_s-site could be identified. The possible I_s-site on the GGDEF domain, which was seen in the YdeH^{GGDEF} structure needs further investigation by using the single mutants of this site. In addition, the I_s-site on the GGDEF domain, which is used in PleD, WspR and in the DGC from *M. aquaeolei* corresponding to residue Arg152 in YdeH, is also a possible site, which has to be analyzed. As it is the case for PleD,

where two different I_s-site are present, which operate redundantly, the integrity of only one of the I_s-sites is required for noncompetitive product inhibition. Therefore a correct I_s-site might already be tested, but it showed no effect on the inhibition by c-di-GMP, because another I_s-site is existing in parallel and still inhibits YdeH.

3.1.11 Model of zinc dependent regulation of YdeH activity

The presented structures of full-length YdeH and YdeH^{CZB} represent an inhibited state of YdeH, because they have zinc bound to the CZB domain. This is in agreement that in the full-length structure the two substrates molecules are too far away to react with each other and produce c-di-GMP. The binding of zinc to the CZB domains keeps the helices of each CZB domain in a compact and fixed arrangement, in which the linker between the CZB and GGDEF domain is defined, too. Thereby the two GGDEF domains are fixed in an incompetent state, which does not allow the encounter of the two GGDEF domains to form c-di-GMP.

Upon zinc release, the variability of the CZB domain is increased, because the helices α1, α2 and α3 are not connected to each other anymore via the zinc. One possibility to activate YdeH after zinc release, could be due to the rearrangement of the helices in the CZB domain, which would bring both helices α4 and consequently also both GGDEF domains into a proximity, where the catalysis can take place (Figure 3.49). Thus the inhibition of YdeH by zinc can be explained by the fixation of the inter domain region.

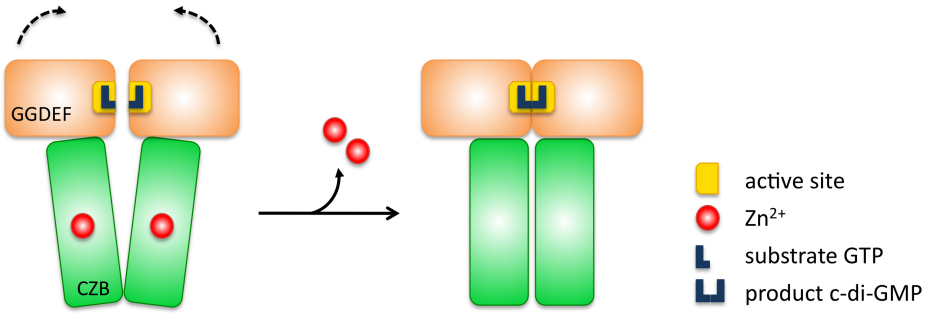


Figure 3.49: Mechanistic model of YdeH regulation by zinc. The binding of zinc to the CZB domain fixes YdeH in a conformation, where the two actives sites of the GGDEF domain can not approach each other to perform catalysis. If the zinc is released, the variability of the CZB domain is increased, which allows the GGDEF domains to form a competent dimer.

3.2 Summary and Outlook

The synthesis of the second messenger c-di-GMP is performed by diguanylate cyclases. They get activated either by external or internal signals, which interact with N-terminal sensory domains of the DGCs. Only very few of these signals are discovered so far. In order to investigate new input signals and to study the regulation of the diguanylate cyclases, the DGC YdeH from *E. coli* was chosen for structural and biochemical studies.
In this thesis, the three-dimensional structure of YdeH could be solved, which provided new insights in the input signal and regulation of YdeH. The N-terminal sensory domain of YdeH shows a new fold, a four helical bundle, which harbors a zinc-binding site compromised of three histidines and one cysteine. It could be shown using enzymatical studies, that this N-terminal sensory domain is necessary for the DGC activity of YdeH. This domain mediates the dimerization of YdeH, which is essential for the catalysis since two monomers built up the active site. It could be shown that YdeH is a permanent dimer. This raises the question, if YdeH is permanently active or how it is regulated independent from the dimerization. It could be shown that the DGC activity of YdeH is regulated by zinc via binding to the N-terminal sensory domain. YdeH has a very high affinity for zinc in the femtomolar range, which is comparable to other zinc sensors. Since the free cellular zinc concentration is very low and the affinity of the zinc binding proteins is very high, there are indications that mechanisms exists by which the zinc is directly passed from protein to protein depending on the relative affinities. The unstructured loop region close to the zinc binding cysteine in the N-terminal domain is ideally suited for this fast zinc exchange. YdeH represents the first example of a biological zinc-sensor that exerts its downstream effects posttranscriptionally and the first example of a metal sensory c-di-GMP signaling protein.
In the structure of full-length YdeH, substrate binding to the active site could be shown, in which the ribose and the α phosphate moiety are not involved in protein interactions. The dimer arrangement of the two monomers is not in a competent conformation, because the substrate molecules are too far apart from each other to perform catalysis. This is not surprising, since the structure is in an inhibited state because zinc is bound to the N-terminal domain.
Product binding to the active site could also be shown in the structure of $YdeH^{GGDEF}$. Here, c-di-GMP crosslinks the two active sites. The guanyl bases of c-di-GMP bind to the same pocket as the substrate, but the ribose and the α phosphate moiety adapt different conformations and are involved in protein interactions. In this structure, the dimer arrangement of the two DGC domains is stabilized via two salt bridges, which were shown to be essential for the DGC activity of YdeH.
To get insights into the competent dimer arrangement of YdeH, a model was generated

assuming an inline attack. In this model however, the N-termini of both DGC domains are too far apart from each other to allow dimerization via the N-terminal CZB domains. Although the competent state of YdeH is not known, the transition from the substrate bound state to the product bound state is attended by reorientation of the DGC domains relative to each other. The inhibition of zinc by binding to the N-terminal sensory domain can be explained by the fixation of the inter domain region. This fixed position prevents the variability needed to perform catalysis.
In the presented structures of YdeH, dimeric c-di-GMP could be identified, which was bound to the inhibitory site of YdeH. In enzymatic experiments, product inhibition could be detected, however the inhibition constant was far above the physiological c-di-GMP concentration. This low affinity could be caused by the variation in the I-site motif compared to PleD or WspR. A secondary I-site, which is visible in the YdeH^{GGDEF} structure could not be assigned definitely so far.

Product inhibition of YdeH will have to be investigated in further experiments by studying the single mutations of the observed possible secondary I-site in the YdeH^{GGDEF} structure. In addition, the secondary I-site, which is seen in PleD and WspR has to be explored. One important goal would be to determine the structure of a competent state of YdeH without zinc bound to the N-terminal CZB domain. This would provide insights into the catalytic mechanism and particularly into the zinc-mediated regulation of YdeH.

4 Large scale enzymatic production of c-di-GMP by YdeH

4.1 Publication: Efficient enzymatic production of the bacterial second messenger c-di-GMP by the diguanylate cyclase YdeH from *E. coli*

Appl Biochem Biotechnol (2011) 163:71–79
DOI 10.1007/s12010-010-9017-x

Efficient Enzymatic Production of the Bacterial Second Messenger c-di-GMP by the Diguanylate Cyclase YdeH from *E. coli*

Franziska Zähringer · Claudia Massa · Tilman Schirmer

Received: 19 January 2010 / Accepted: 15 June 2010 /
Published online: 29 June 2010

Abstract Cyclic di-GMP (c-di-GMP) is an almost universal bacterial second messenger involved in the regulation of cell surface-associated traits and the persistence of infections. GGDEF and EAL domain-containing proteins catalyse c-di-GMP synthesis and degradation, respectively. We report the enzymatic large-scale synthesis of c-di-GMP, making use of the GGDEF domain-containing protein YdeH from *Escherichia coli*. Overexpression and purification of YdeH have been established, and the conditions for c-di-GMP synthesis were optimised. In contrast to the chemical synthesis of c-di-GMP, enzymatic c-di-GMP production is a one-step reaction that can easily be performed with the equipment of a standard biochemical lab. The protocol allows the production of milligram amounts of c-di-GMP within 1 day and paves the way for extensive biochemical and biophysical studies on c-di-GMP-mediated processes.

Keywords c-di-GMP · Diguanylate cyclase · GGDEF domain · Enzymatic synthesis · *E. coli*

Introduction

Bis-(3′-5′)-cyclic di-GMP (c-di-GMP) is an important bacterial second messenger involved in the regulation of a number of complex physiological processes. C-di-GMP was first identified 20 years ago as an activating factor of cellulose synthase in *Acetobacter xylinum* [1], but in the mean time, it has been shown to play a central role, among others, in the transition between a motile, single-cell state to a sessile, surface-attached state found in biofilms [2–5]. Its relevance for the virulence of pathogenic bacteria is well established [2–6].

C-di-GMP is synthesised by the condensation of two GTP molecules. This reaction is catalysed by the GGDEF domain of diguanylate cyclases (DGCs) [7–10], whereas EAL [7, 11–13] and HD-GYP [14] domains hydrolyse the compound to yield the linear pGpG

F. Zähringer · C. Massa · T. Schirmer (✉)
Core Program Structural Biology and Biophysics, Biozentrum, University of Basel,
Klingelbergstrasse 70, CH-4056 Basel, Switzerland
e-mail: tilman.schirmer@unibas.ch

dinucleotide. All these catalytic domains are typically found in combinations with other, mostly sensory or regulatory, domains and control the cellular level of c-di-GMP.

Much about the mechanisms of c-di-GMP signalling in bacteria has still to be discovered, in particular since only a few targets of c-di-GMP have been identified. Characterised c-di-GMP receptors comprise PilZ domains [15] and degenerated GGDEF and EAL domains [16, 17]. Furthermore, the GEMM riboswitches have recently been discovered to regulate gene expression via c-di-GMP binding [18–20].

Recent studies have demonstrated that exogenous c-di-GMP treatment inhibits adhesive *Staphylococcus aureus* cell-to-cell interactions and biofilm formation [21] and that it is effective also in a mouse model [22]. These findings make c-di-GMP an interesting candidate as a potential antimicrobial agent. It was demonstrated that c-di-GMP stimulates the immune system to prevent bacterial infections and is therefore evaluated as a potential vaccine adjuvant candidate [23–25]. It was also reported that c-di-GMP inhibits cancer cell proliferation in vitro and could therefore be used as a therapeutic agent [26].

In order to carry out extensive studies on c-di-GMP-mediated processes, sufficient supply of this compound is crucial. The reported chemical syntheses of c-di-GMP are multistep reactions, which make them time-consuming, expensive and inefficient [27–33]. Most of them make use of phosphotriester, phosphoamidite or H-phosphonate chemistry based on air- or water-sensitive reagents and involve several chromatographic purification steps.

In contrast, DGC-catalysed synthesis of c-di-GMP from GTP appears straightforward, as has been reported for PleD, VCA0956 and WspR [12, 13, 34–37]. However, all these DGCs show potent allosteric product inhibition with K_i values in the low micromolar range. Furthermore, the purified proteins are rather unstable in solution. Recently, a DGC from a thermophilic organism (*Thermotoga maritima*) has been employed for the enzymatic synthesis of c-di-GMP [38]. A fragment comprising only its DGC domain with the allosteric inhibition site (I-site) mutated to impede product inhibition has been exploited. Large amounts (about 20 mg) of c-di-GMP were obtained employing 1 mg of this modified version of the enzyme and by replenishing repeatedly the reaction mixture with the substrate GTP.

In this paper, we report a large-scale production procedure of c-di-GMP employing, as an alternative enzyme, YdeH from *Escherichia coli*. This enzyme consists of an N-terminal domain of unknown fold and a C-terminal GGDEF domain. We have shown previously [39] that YdeH, in contrast to most of the other DGCs, is a constitutively active dimer and exhibits a turnover of 1.6 min^{-1} and a K_m of 17 μM. The YdeH catalysed reaction shows non-competitive product inhibition only at comparatively large c-di-GMP concentration (K_i=44 μM), in contrast to the aforementioned DGCs. Furthermore, even at very high c-di-GMP concentration (>1 mM), YdeH shows residual activity (15%) [39]. We show that with this enzyme, complete conversion of GTP to c-di-GMP can be achieved within hours, yielding milligram amount of product.

Material and Methods

Expression and Purification of YdeH

C-terminally His6-tagged YdeH was expressed in the pET28b plasmid in the *E. coli* Rosetta strain. The transformed cells were cultivated at 37 °C in Luria–Bertani medium supplemented with ampicillin (100 μg/mL) and chloramphenicol (34 μg/mL) until the OD_{600} had reached 0.7. Gene expression was induced by adding 1 mM isopropyl-β-D-thiogalactopyranoside

(IPTG). After 4 h of incubation, the cells were harvested by centrifugation (6,800×*g*, 10 min, 4 °C) and the pellets were frozen at −20 °C.

Prior to lysis, the frozen cell pellets were thawed and resuspended in Ni-A buffer (50 mM NaH_2PO_4, pH 7.5, 200 mM NaCl, 10 mM imidazole, 50 mM L-glutamic acid, 50 mM L-arginine). After the addition of DNAse (2.5 μg/mL; Sigma) and EDTA-free protease-inhibitor cocktail (Roche, 1 tablet/50 mL buffer), the cells were disrupted with a French press (Thermo Spectronic) at 15,000 psi. The lysate was cleared by centrifugation (28,000×*g*, 45 min, 4 °C), and the supernatant was filtered (0.22 μm) and loaded onto a 5-mL HisTrap column (GE Healthcare). After washing the column with Ni-A buffer (10 column volumes (CV)), the protein was eluted with a linear gradient of imidazole from 10 to 500 mM in 10 CV. The pooled fractions were concentrated to 1 mL and further purified by size-exclusion chromatography using a Superdex 75 16/60 column (GE Healthcare) and SEC buffer (20 mM Tris–HCl, pH 7.6, 150 mM NaCl, 50 mM L-glutamic acid, 50 mM L-arginine). The chromatographic runs were carried out with an Äkta Purifier FPLC unit and monitored with Unicorn software. The purification process was monitored by SDS-PAGE followed by Coomassie blue R-250 staining. Protein concentration was determined by measuring the absorbance at 280 nm (ε_{280}=39,880 $M^{-1}cm^{-1}$).

Enzyme Assay

The initial rate of c-di-GMP synthesis was measured by following the substrate consumption using ion exchange chromatography. The reaction mixture (100 μL) contained 2 μM purified YdeH, 100 μM GTP (Sigma), and 5 mM $MgCl_2$. The effect of pH on the activity of YdeH was tested in 50 mM MES, pH 6.5–7.5, 50 mM Tris–HCl, pH 7.5–8.5, and 50 mM CHES, pH 9.0–10.0. Also, the effect of different salts (LiCl, NaCl and KCl) and NaCl concentrations (25 mM–1.5 M) was investigated. The reaction was stopped by heating of the sample for 2 min at 99 °C. Subsequently, 100 μL of the reaction mixture was diluted in 900 μL 5 mM NH_4HCO_3, pH 8.0, filtered (0.22 μm) and loaded on an ion-exchange column (ResourceQ 1 mL, GE Healthcare). The nucleotides were separated with a gradient from 0.005 to 1 M NH_4HCO_3, pH 8.0, in 14 CV. The amount of substrate and reaction product was determined by integration of the UV absorption (253 nm) peaks. The procedure was calibrated with GTP (Sigma) and c-di-GMP (Biolog, Bremen) standards of known concentrations.

Enzymatic Production of c-di-GMP

For the synthesis of c-di-GMP, 2 μM YdeH was incubated with 500 μM GTP in 50 mM Tris–HCl, pH 7.5, 50 mM NaCl and 5 mM $MgCl_2$ for 5 h. The progress of the reaction was followed by withdrawing samples (100 μL) that were inactivated by heat denaturation. Subsequently, the enzyme was removed by filtering and the sample content analysed by ion-exchange chromatography. After completion of the reaction, the mixture was heated for 5 min at 99 °C, centrifuged (4,500×*g*, 10 min, room temperature) and filtered (0.22 μm). For purification of the reaction product, Et_3NHCO_3 to a final concentration of 5 mM was added to the reaction mixture and loaded on a reversed-phase chromatography column (1.7 mL per 1 mL resin Resource RPC, GE Healthcare). After washing the column with 5 mM Et_3NHCO_3 (10 CV), c-di-GMP was eluted with a linear gradient of ethanol from 0% to 50% in 10 CV. The reversed-phase chromatography was carried out with an Äkta Purifier FPLC unit and monitored with Unicorn software. The c-di-GMP-containing fractions were pooled and lyophilised. The powder was dissolved in water and the concentration was

determined by measuring the optical density at 253 nm of an aliquot diluted to 10 μM (assuming an ε_{253} of 23,700 $M^{-1}cm^{-1}$ as reported in [29]).

Analytical Methods

Purified c-di-GMP was analysed by high-performance liquid chromatography coupled to mass spectrometry. The samples were diluted 50-fold with 10 mM NH_4OAc, pH 6.4, and injected into a Supercosil™ LC-18-T column (Supelco). The run was performed at a flow rate of 0.7 mL/min using a linear gradient (0–50%) of acetonitrile. The electron spray ionisation mass spectrometry with a time-of-flight analyser was performed using a microTOF Focus system (Bruker Daltronics). The capillary voltage was 4500 V, and the end-plate offset was 500 V (negative mode); the dry temperature was 200 °C, the dry gas flow was 9 L/min, and the nebulizer pressure was 2 bar.

To test the purity of the obtained c-di-GMP, a sample was analysed by ^{1}H-nuclear magnetic resonance (NMR) spectroscopy (Bruker Advance 600) in 10 mM Tris–HCl (D11), pH 7.4 (Cambridge Isotope Laboratories, Inc.), 150 mM NaCl, 50 μM EDTA at 298 K. For comparison, spectra of synthetic c-di-GMP (Biolog) and GTP (Sigma) were acquired.

Results and Discussion

Expression and Purification of YdeH

YdeH expression was tested in different *E. coli* strains and at different temperatures. In all conditions, most of the expressed YdeH was found in inclusion bodies. The highest amount of soluble protein was achieved in the Rosetta strain at 37 °C, 3–4 h after IPTG induction. YdeH was purified to homogeneity using a two-step purification procedure, consisting of Ni-affinity and size-exclusion chromatography. To achieve a protein concentration above 0.8 mg/mL, the addition of arginine and glutamic acid to the buffers was essential to avoid protein aggregation [40]. The obtained protein was pure and no degradation fragments were present, as demonstrated by the occurrence of a single band on SDS-PAGE (Fig. 1). The final yield was about 5 mgL^{-1} culture.

Optimisation and Characterisation of the Enzymatic Reaction

Substrate and product amounts were analysed by ion-exchange chromatography. A peak at a somewhat smaller elution volume than that of the final product was observed in the

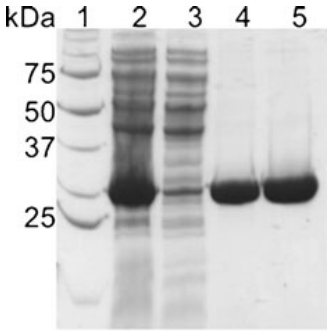


Fig. 1 SDS-PAGE of the purification of YdeH. A 12% (*w*/*v*) gel loaded with: *lane 1* molecular weight marker, *lane 2* cell extract, *lane 3* soluble fraction of the cell extract, *lane 4* purified YdeH after Ni column, *lane 5* purified YdeH after size-exclusion chromatography

Appl Biochem Biotechnol (2011) 163:71–79

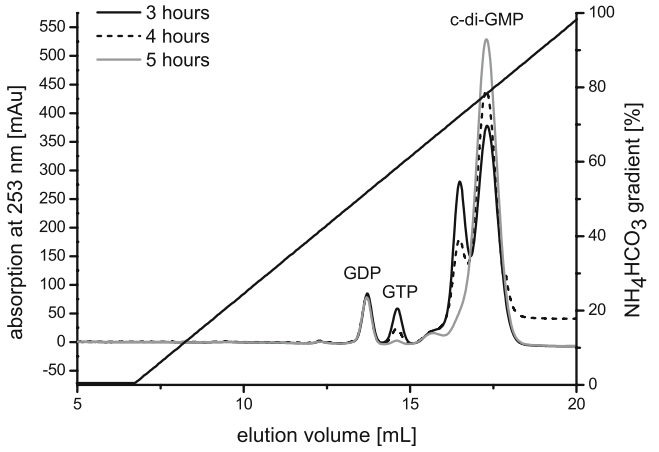


Fig. 2 Time course of the synthesis of c-di-GMP. The samples were analysed using ion-exchange chromatography. After 5 h, the reaction is virtually complete. The leftmost small and constant peak corresponds to a GDP contamination in the commercial GTP batch

chromatograms taken at early to intermediate time points, but was absent after completion of the reaction (Fig. 2). As demonstrated by MS, this transient peak did not correspond to a reaction intermediate, but had the mass of c-di-GMP. Probably, the two peaks correspond to different oligomeric states of c-di-GMP.

The optimum pH for the synthesis of c-di-GMP was found to be at pH 7.5 in Tris buffer (Fig. 3). Variation of the salt (LiCl, NaCl, KCl) had only a marginal effect. The enzymatic activity was virtually the same in 50 mM NaCl and 50 mM KCl, but was found to be reduced by about 20% in 50 mM LiCl. Similarly, variation of the NaCl concentration (25 mM–1.5 M) had no significant effect on activity. Therefore, we chose a comparatively low NaCl concentration (50 mM) to be compatible with the requirements for the subsequent c-di-GMP purification procedure.

Enzymatic c-di-GMP Production and Purification

For the large-scale synthesis of c-di-GMP, the molar ratio of enzyme to substrate has been optimised in order to obtain complete conversion within a convenient period. Using 2 μM YdeH and 500 μM GTP, virtually complete substrate conversion was achieved within 5 h,

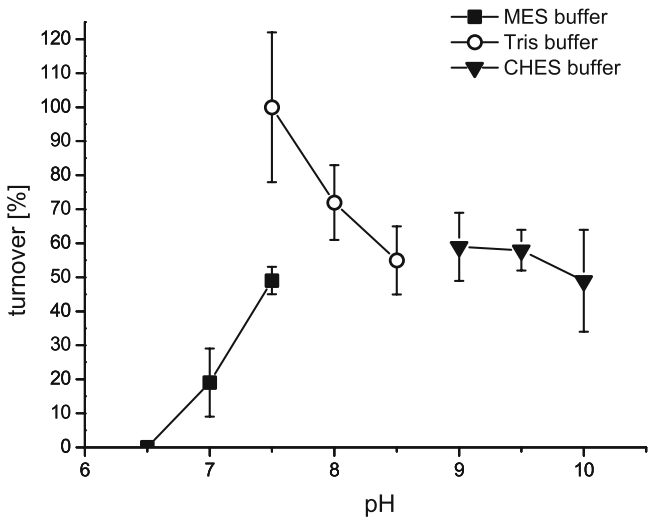


Fig. 3 pH dependence of the enzymatic activity of YdeH. The optimum reaction velocity for the synthesis of c-di-GMP was found in 50 mM Tris–HCl, pH 7.5

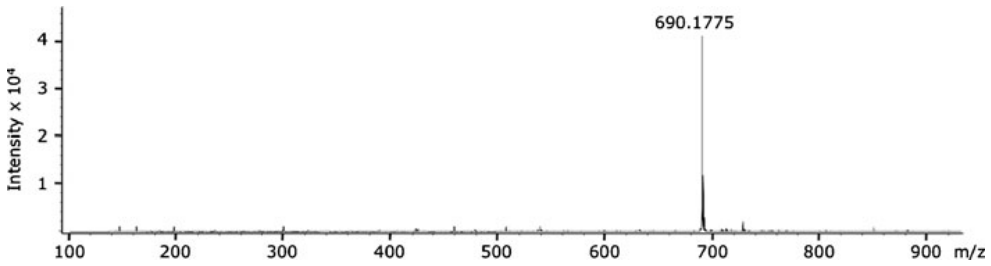


Fig. 4 Convoluted mass spectrum of enzymatically produced c-di-GMP. The measured mass of 690.1775 g/mol is in agreement with the calculated mass of 690.0870 g/mol

as shown in Fig. 2. Initially, the reaction volume was 0.1 mL, but was eventually upscaled to 0.5 L without any detrimental effect. C-di-GMP was purified on a reversed-phase column with ethanol as eluent. The Resource RPC resin was chosen because of its stability at basic pH values. The mobile phase was adjusted so as to prevent binding of the mono-nucleotides GDP and GTP. Ethanol has the advantage of being non-toxic compared to other typically used eluents such as acetonitrile or methanol. The procedure yielded 75 mg of pure c-di-GMP employing 30 mg of YdeH. Considering the ease of YdeH purification, further optimisation of the procedure by recycling of the enzyme as in [38] was not considered.

Product Analysis

LC/MS clearly identified the reaction product as c-di-GMP (Fig. 4). No impurities were detected in the chromatogram. The observed mass of c-di-GMP was in perfect agreement with the calculated mass. In addition, the ^{1}H-NMR spectra of purified c-di-GMP completely coincided with the spectrum of chemically synthesised c-di-GMP and showed no traces of the substrate GTP (Fig. 5).

Conclusions

Here, we have described an alternative way for the enzymatic synthesis of the bacterial second messenger c-di-GMP. We use the DGC YdeH from *E. coli*, which is an ideal

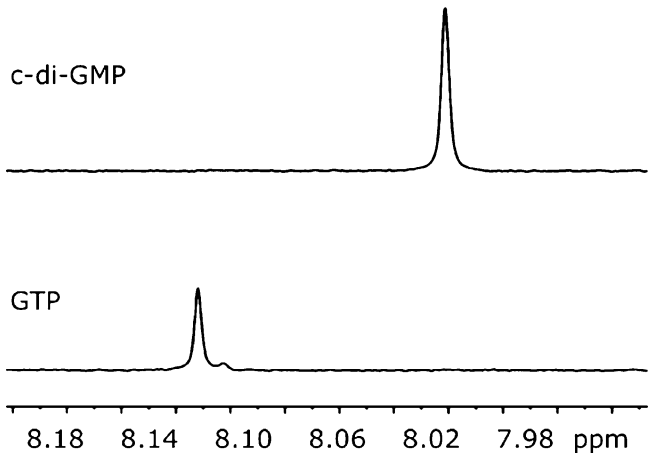


Fig. 5 Part of the ^{1}H-NMR spectrum (600 MHz) showing the chemical shift of H8 of enzymatically produced c-di-GMP (*top*, δ=8.04) and of GTP (*bottom*, δ=8.13) as reference. Comparison demonstrates that the c-di-GMP sample was free of GTP. Chemical shifts are relative to H_2O (δ=4.77, 298 K)

candidate because it is constitutively active in vitro and exhibits only weak product inhibition; thus, c-di-GMP concentrations of up to 0.25 mM can be achieved. Our system allows producing 2.5 mg of pure c-di-GMP per milligram of YdeH with standard biochemical lab equipment.

Enzymatic c-di-GMP production outperforms chemical synthesis with respect to time, costs and number of chromatographic purification steps [27–33]. Moreover, most of the chemical syntheses are based on air- and water-sensitive reagents. The enzymatic procedures described in the literature [12, 13, 34–37], however, employ rather unstable DGCs that, in addition, show strong product inhibition and therefore do not yield complete substrate conversion. The recently described procedure using a DGC from a thermophilic organism (*T. maritima*) [38] solved the solubility issue and provides the best specific yield so far (20 mg/mg enzyme), but had to be mutated in the allosteric inhibition site (I-site) to abolish product inhibition.

Acknowledgement We thank Dr. Alexander Böhm, Biozentrum Basel, for the gift of the plasmid, Dr. Martin Allan, Biozentrum Basel, for help with ^{1}H-NMR spectroscopy and Dietrich Samoray for suggesting to test the effect of osmolytes on YdeH solubility. The work was supported by Swiss National Science Foundation grant 3100A0-105587.

References

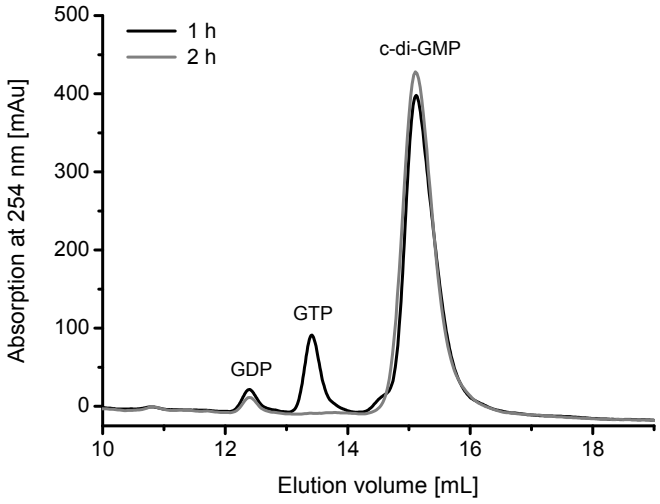
1. Ross, P., Weinhouse, H., Aloni, Y., Michaeli, D., Weinberger-Ohana, P., Mayer, R., et al. (1987). Regulation of cellulose synthesis in *Acetobacter xylinum* by cyclic diguanylic acid. *Nature, 325*, 279–381.
2. Jenal, U., & Malone, J. (2006). Mechanisms of cyclic-di-GMP signaling in bacteria. *Annual Review of Genetics, 40*, 385–407.
3. Tamayo, R., Pratt, J. T., & Camilli, A. (2007). Roles of cyclic diguanylate in the regulation of bacterial pathogenesis. *Annual Review of Microbiology, 61*, 131–148.
4. Schirmer, T., & Jenal, U. (2009). Structural and mechanistic determinants of c-di-GMP signaling. *Nature Reviews Microbiology, 7*, 724–735.
5. Pesavento, C., & Hengge, R. (2009). Bacterial nucleotide-based second messengers. *Current Opinion in Microbiology, 12*, 170–176.
6. Tamayo, R., Schild, S., Pratt, J. T., & Camilli, A. (2008). Role of cyclic di-GMP during el tor biotype *Vibrio cholerae* infection: Characterization of the in vivo-induced cyclic di-GMP phosphodiesterase CdpA. *Infection and Immunity, 76*, 1617–1627.
7. Tal, R., Wong, H. C., Calhoon, R., Gelfand, D., Fear, A. L., Volman, G., et al. (1998). Three cdg operons control cellular turnover of cyclic di-GMP in *Acetobacter xylinum*: Genetic organization and occurrence of conserved domains in isoenzymes. *Journal of Bacteriology, 80*, 4416–4425.
8. Chan, C., Paul, R., Samoray, D., Amiot, N. C., Giese, B., Jenal, U., et al. (2004). Structural basis of activity and allosteric control of diguanylate cyclase. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 17084–17089.
9. Paul, R., Weiser, S., Amiot, N. C., Chan, C., Schirmer, T., Giese, B., et al. (2004). Cell cycle-dependent dynamic localization of a bacterial response regulator with a novel di-guanylate cyclase output domain. *Genes & Development, 18*, 715–727.
10. Ryjenkov, D. A., Tarutina, M., Moskvin, O. V., & Gomelsky, M. (2005). Cyclic diguanylate is a ubiquitous signaling molecule in bacteria: Insights into biochemistry of the GGDEF protein domain. *Journal of Bacteriology, 187*, 1792–1798.
11. Schmidt, A. J., Ryjenkov, D. A., & Gomelsky, M. (2005). The ubiquitous protein domain EAL is a cyclic diguanylate-specific phosphodiesterase: Enzymatically active and inactive EAL domains. *Journal of Bacteriology, 187*, 4774–4781.
12. Christen, M., Christen, B., Folcher, M., Schauerte, A., & Jenal, U. (2005). Identification and characterization of a cyclic di-GMP-specific phosphodiesterase and its allosteric control by GTP. *The Journal of Biological Chemistry, 280*, 30829–30837.

13. Rao, F., Yang, Y., Qi, Y., & Liang, Z. X. (2008). Catalytic mechanism of cyclic di-GMP-specific phosphodiesterase: A study of the EAL domain-containing RocR from *Pseudomonas aeruginosa*. *Journal of Bacteriology, 190*, 3622–3631.
14. Galperin, M. Y. (2006). Structural classification of bacterial response regulators: Diversity of output domains and domain combinations. *Journal of Bacteriology, 188*, 4169–4182.
15. Amikam, D., & Galperin, M. Y. (2006). PilZ domain is part of the bacterial c-di-GMP binding protein. *Bioinformatics, 22*, 3–6.
16. Duerig, A., Abel, S., Folcher, M., Nicollier, M., Schwede, T., Amiot, N., et al. (2009). Second messenger-mediated spatiotemporal control of protein degradation regulates bacterial cell cycle progression. *Genes & Development, 23*, 93–104.
17. Tschowri, N., Busse, S., & Hengge, R. (2009). The BLUF-EAL protein YcgF acts as a direct anti-repressor in a blue-light response of *Escherichia coli*. *Genes & Development, 15*, 522–534.
18. Sudarsan, N., Lee, E. R., Weinberg, Z., Moy, R. H., Kim, J. N., Link, K. H., et al. (2008). Riboswitches in eubacteria sense the second messenger cyclic di-GMP. *Science, 321*, 411–413.
19. Smith, K. D., Lipchock, S. V., Ames, T. D., Wang, J., Breaker, R. R., & Strobel, S. A. (2009). Structural basis of ligand binding by a c-di-GMP riboswitch. *Nature Structural & Molecular Biology, 16*, 1218–1223.
20. Kulshina, N., Baird, N. J., & Ferré-D'Amaré, A. R. (2009). Recognition of the bacterial second messenger cyclic diguanylate by its cognate riboswitch. *Nature Structural & Molecular Biology, 16*, 1212–1217.
21. Karaolis, D. K., Rashid, M. H., Chythanya, R., Luo, W., Hyodo, M., & Hayakawa, Y. (2005). C-di-GMP (3′-5′-cyclic diguanylic acid) inhibits *Staphylococcus aureus* cell–cell interactions and biofilm formation. *Antimicrobial Agents and Chemotherapy, 49*, 1029–1038.
22. Brouillette, E., Hyodo, M., Hayakawa, Y., Karaolis, D. K., & Malouin, F. (2005). 3′-5′-Cyclic diguanylic acid reduces the virulence of biofilm-forming *Staphylococcus aureus* strains in a mouse model of mastitis infection. *Antimicrobial Agents and Chemotherapy, 49*, 3109–3113.
23. Karaolis, D. K., Means, T. K., Yang, D., Takahashi, M., Yoshimura, T., Muraille, E., et al. (2007). Bacterial c-di-GMP is an immunostimulatory molecule. *Journal of Immunology, 178*, 2171–2181.
24. Karaolis, D. K., Newstead, M. W., Zeng, X., Hyodo, M., Hayakawa, Y., Bhan, U., et al. (2007). Cyclic di-GMP stimulates protective innate immunity in bacterial pneumonia. *Infection and Immunity, 75*, 4942–5450.
25. Ogunniyi, A. D., Paton, J. C., Kirby, A. C., McCullers, J. A., Cook, J., Hyodo, M., et al. (2008). C-di-GMP is an effective immunomodulator and vaccine adjuvant against pneumococcal infection. *Vaccine, 26*, 4676–4685.
26. Karaolis, D. K., Cheng, K., Lipsky, M., Elnabawi, A., Catalano, J., Hyodo, M., et al. (2005). 3′,5′-Cyclic diguanylic acid (c-di-GMP) inhibits basal and growth factor-stimulated human colon cancer cell proliferation. *Biochemical and Biophysical Research Communications, 329*, 40–45.
27. Ross, P., Mayer, R., Weinhouse, H., Amikam, D., Huggirat, Y., Benziman, M., et al. (1990). The cyclic diguanylic acid regulatory system of cellulose synthesis in *Acetobacter xylinum*. Chemical synthesis and biological activity of cyclic nucleotide dimer, trimer, and phosphothioate derivatives. *The Journal of Biological Chemistry, 265*, 18933–18943.
28. Kawai, R., Nagata, R., Hirata, A., & Hayakawa, Y. (2003). A new synthetic approach to cyclic bis (3′→5′)diguanylic acid. *Nucleic Acids Symposium Series, 3*, 103–104.
29. Hayakawa, Y., Nagata, R., Hirata, A., Hyodo, M., & Kawai, R. (2003). A facile synthesis of cyclic bis (3′→5′)diguanylic acid. *Tetrahedron, 59*, 6465–6471.
30. Zhang, Z., Gaffney, B. L., & Jones, R. A. (2004). C-di-GMP displays a monovalent metal ion-dependent polymorphism. *Journal of the American Chemical Society, 26*, 16700–16701.
31. Amiot, N., Heintz, K., & Giese, B. (2006). New approach for the synthesis of c-di-GMP and its analogues. *Synthesis, 24*, 4230–4236.
32. Yan, H., & Aguilar, A. L. (2007). Synthesis of 3′,5′-cyclic diguanylic acid (cdiGMP) using 1-(4-chlorophenyl)-4-ethoxypiperidin-4-yl as a protecting group for 2′-hydroxy functions of ribonucleosides. *Nucleosides, Nucleotides & Nucleic Acids, 26*, 189–204.
33. Kiburu, I., Shurer, A., Yan, L., & Sintim, H. O. (2008). A simple solid-phase synthesis of the ubiquitous bacterial signaling molecule, c-di-GMP and analogues. *Molecular Biosystems, 4*, 518–520.
34. Tamayo, R., Tischler, A. D., & Camilli, A. (2005). The EAL domain protein VieA is a cyclic diguanylate phosphodiesterase. *The Journal of Biological Chemistry, 280*, 33324–33330.
35. Kazmierczak, B. I., Lebron, M. B., & Murray, T. S. (2006). Analysis of FimX, a phosphodiesterase that governs twitching motility in *Pseudomonas aeruginosa*. *Molecular Microbiology, 60*, 1026–1043.
36. Merighi, M., Lee, V. T., Hyodo, M., Hayakawa, Y., & Lory, S. (2007). The second messenger bis-(3′-5′)-cyclic-GMP and its PilZ domain-containing receptor Alg44 are required for alginate biosynthesis in *Pseudomonas aeruginosa*. *Molecular Microbiology, 65*, 876–895.

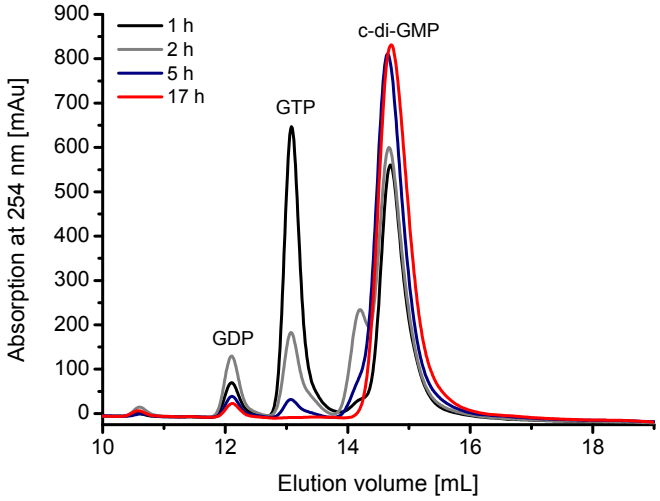
37. Hickman, J. W., & Harwood, C. S. (2008). Identification of FleQ from *Pseudomonas aeruginosa* as a c-di-GMP-responsive transcription factor. *Molecular Microbiology, 69*, 376–389.
38. Rao, F., Pasunooti, S., Ng, Y., Zhuo, W., Lim, L., Liu, A. W., et al. (2009). Enzymatic synthesis of c-di-GMP using a thermophilic diguanylate cyclase. *Analytical Biochemistry, 389*, 138–142.
39. Böhm, A., Steiner, S., Zähringer, F., Casanova, A., Hamburger, F., Ritz, D., et al. (2009). Second messenger signalling governs *Escherichia coli* biofilm induction upon ribosomal stress. *Molecular Microbiology, 72*, 1500–1516.
40. Golovanov, A. P., Hautbergue, G. M., Wilson, S. A., & Lian, L. Y. (2004). A simple method for improving protein solubility and long-term stability. *Journal of the American Chemical Society, 126*, 8933–8939.

4.2 Improved protocol

The protocol in the publication 'Efficient enzymatic production of the bacterial second messenger c-di-GMP by the diguanylate cyclase YdeH from *E. coli*' was developed before it was known that YdeH is inhibited by zinc. To optimize this protocol, zinc-free YdeH was used. By using 2 μM zinc-free YdeH in the same conditions as described in the publication (500 μM GTP in 50 mM Tris-HCl, pH 8.0, 50 mM NaCl and 5 mM $MgCl_2$), the reaction was complete after 2 h (Figure 4.1). This is 2.5 times faster as using YdeH that is not zinc-free. In order to achieve a higher yield of c-di-GMP the GTP concentration was doubled to 1 mM. In Figure 4.1 the time course of this reaction is shown. After 5 h, most of the GTP is consumed and after 17 h the reaction is complete. With this new protocol, the yield could be doubled, which results in 5 mg of pure c-di-GMP per milligram of YdeH.



(A)



(B)

Figure 4.1: Time course of the synthesis of c-di-GMP. The samples were analyzed using ion-exchange chromatography. (A) The reaction was performed by using 2 μM zinc-free YdeH and 500 μM GTP. After 2 h, the reaction is virtually complete. (B) The reaction was performed by using 2 μM zinc-free YdeH and 1 mM GTP. After 5 h, the reaction is nearly complete. The leftmost small peak corresponds to a GDP contamination in the commercial GTP batch.

5 Characterization of PgaA and PgaB, members of the c-di-GMP controlled exopolysaccharide synthesis machinery

5.1 Results and Discussion

5.1.1 Design and cloning of PgaA and PgaB constructs

In order to perform structural and biochemical studies on PgaA from *E. coli*, several constructs were designed. Various expression vectors were chosen for the expression of full-length PgaA. PgaA including the N-terminal 32 amino acid containing signal sequence with a C-terminal Histag was cloned in the vectors pET28a and pColdIV. To use other signal sequences for the export to the outer membrane, which allow high expression levels of membrane proteins in *E.coli*, additional expression vectors were selected. Therefore, PgaA was cloned without the signal sequence in the vectors pET22b, pMAL-p5X and pET40b. The signal sequence of PelB from *P. aeruginosa* and a C-terminal Histag are included in the pET22b vector. In the pMAL-p5X vector, PgaA is fused to maltose-binding protein (MBP) using its signal sequence. This construct has a cleavage site for the protease Factor Ax between MBP and PgaA. In the pET40b vector, PgaA is fused to the periplasmic protein DsbC, which includes a C-terminal Histag and a cleavage site for thrombin between DsbC and PgaA.
To study the soluble domain of PgaA, a construct containing the N-terminal domain with a C-terminal Histag, was cloned without the signal sequence in the pET28a vector. All constructs were cloned as described in section 2.1 and are listed in Table 5.1.

For the expression of PgaB from *E. coli*, two different construct were designed. Both constructs started with amino acid 22 after the lipid-attached Cys21, in order to express PgaB in the cytosol. One construct used the expression vector pET21b with an N-terminal Histag, while for the other construct PgaB was cloned with a C-terminal Histag in the pET28a vector. The cloning was performed as described in section 2.1 and the constructs used are listed in Table 5.2.

Table 5.1: PgaA constructs.

Name	PgaA residues	Signal sequence of	N-terminal fusion partner	Protease cleavage site	C-terminal tag
pET28a/PgaA	full-length (33-807)	PgaA	-	-	Histag
pColdIV/PgaA	full-length (33-807)	PgaA	-	-	Histag
pET22b/PgaA	full-length (33-807)	PelB	-	-	Histag
pMal-p5X/PgaA	full-length (33-807)	MBP	MBP	Factor Ax	-
pET40b/PgaA	full-length (33-807)	DsbC	DsbC-Histag	thrombin	-
pET28a/PgaA^{TPR}	soluble domain (33-519)	-	-	-	Histag

Table 5.2: PgaB constructs.

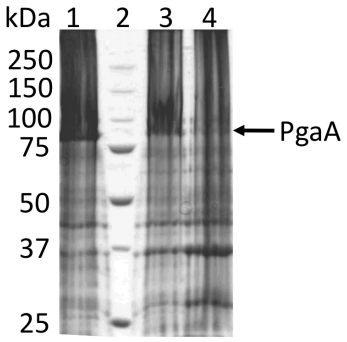
Name	PgaB residues	Features
pET21b/PgaB	22-651	N-terminal Histag
pET28a/PgaB	22-651	C-terminal Histag

5.1.2 Expression and Purification of PgaA

Expression of PgaA

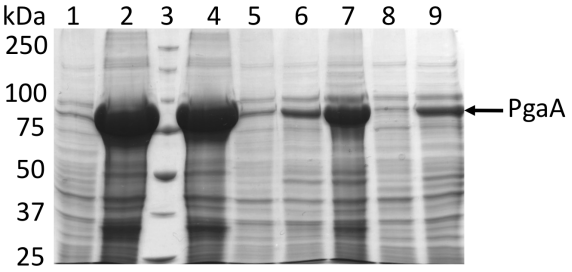
Initial expression tests for pET28a/PgaA were performed in different *E. coli* strains BL21(DE3), BL21(DE3)pLysS, BL21(DE3)pGroEL, Rosetta(DE3), C41(DE3), C43(DE3), ArcticExpress(DE3) and BL21(DE3)omp8/pLysS at two different temperatures of 37 °C and 18 °C. Expression was induced with 1 mM IPTG. For the analysis of the expression test, samples were taken at different time points, from which the membranes were isolated. A detailed description of the procedure is reported in section 2.2.1. In the strains BL21(DE3)pLysS and BL21(DE3)pGroEL PgaA was not expressed. In the other strains, PgaA was highly expressed in inclusion bodies and smaller amounts in the membrane fraction. The best result was achieved in BL21(DE3) at 37 °C after 3 h induction (Figure 5.1).
To further refine the expression conditions, the BL21(DE3) strain was used at 37 °C and 25 °C and the IPTG concentration was varied. For analysis of the expression, the membrane fraction of the taken sample was further separated into inner and outer membrane fractions. Under all tested IPTG concentrations (1 mM, 50 μM and 10 μM) the extracted amount of PgaA from the outer membrane was the same, but the amount of inclusion bodies decreased along with lower IPTG concentrations (Figure 5.2).

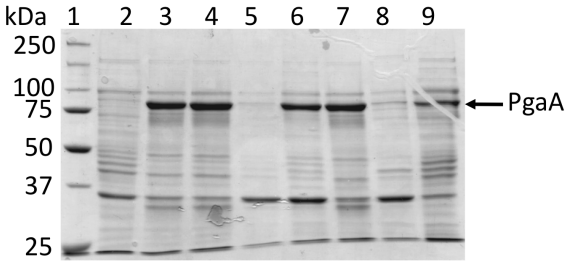
Furthermore, the expression of PgaA was tested using the pColdIV vector. This vector is designed to perform efficient protein expression at low temperature utilizing the promotor derived from the *cspA* gene, which is one of the cold-shock genes in *E. coli*. The different *E.*



lane 1: 3 h after induction, membrane fraction
lane 2: marker
lane 3: 6 h after induction, membrane fraction
lane 4: 20 h after induction, membrane fraction

Figure 5.1: SDS-PAGE of the membrane fractions of the expression test of PgaA in the pET28a vector in BL21(DE3) cells. The expression temperature was 37 ℃ and the induction was carried out with 1 mM IPTG. The best expression was achieved 3 h after induction (lane 1). Sample preparation was done as described in section 2.2.1.





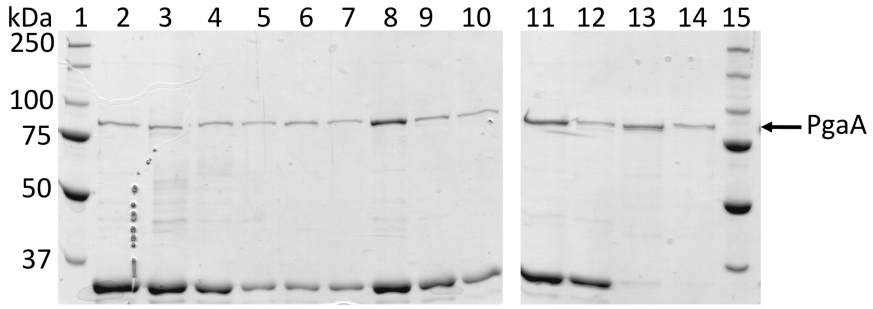
inclusion bodies

lane 1: before induction
lane 2: 37 ℃, 1 mM IPTG, 3 h after induction
lane 3: marker
lane 4: 37 ℃, 50 μM IPTG, 3 h after induction
lane 5: 37 ℃, 10 μM IPTG, 1 h after induction
lane 6: 37 ℃, 10 μM IPTG, 2 h after induction
lane 7: 37 ℃, 10 μM IPTG, 3 h after induction
lane 8: 25 ℃, 10 μM IPTG, 3 h after induction
lane 9: 25 ℃, 10 μM IPTG, 5 h after induction

membrane fractions

lane 1: marker
lane 2: before induction
lane 3: 37 ℃, 1 mM IPTG, 3 h after induction
lane 4: 37 ℃, 50 μM IPTG, 3 h after induction
lane 5: 37 ℃, 10 μM IPTG, 1 h after induction
lane 6: 37 ℃, 10 μM IPTG, 2 h after induction
lane 7: 37 ℃, 10 μM IPTG, 3 h after induction
lane 8: 25 ℃, 10 μM IPTG, 3 h after induction
lane 9: 25 ℃, 10 μM IPTG, 5 h after induction

Figure 5.2: SDS-PAGE of the refined expression test of PgaA in the pET28a vector in BL21(DE3) cells. The IPTG concentration and the expression temperature were varied. The extracted amount of PgaA from the outer membrane is always the same (right panel), but the amount of inclusion bodies is decreased along with lower IPTG concentrations (left panel). Sample preparation was done as described in section 2.2.1.

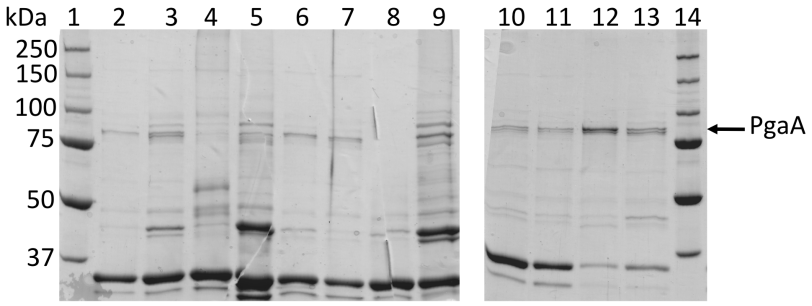


lane 1:	marker	lane 9:	C41(DE3), 18 °C, 20 h
lane 2:	BL21(DE3), 37 °C, 3 h	lane 10:	ArcticExpress(DE3), 18 °C, 20 h
lane 3:	BL21(DE3), 18 °C, 20 h	lane 11:	C43(DE3), 37 °C, 3 h
lane 4:	BL21(DE3)pLysS, 37 °C, 3 h	lane 12:	C43(DE3), 18 °C, 20 h
lane 5:	BL21(DE3)pLysS, 18 °C, 20 h	lane 13:	BL21(DE3)omp8/pLysS, 37 °C, 3 h
lane 6:	Rosetta(DE3), 37 °C, 3 h	lane 14:	BL21(DE3)omp8/pLysS, 18 °C, 20 h
lane 7:	Rosetta(DE3), 18 °C, 20 h	lane 15:	marker
lane 8:	C41(DE3), 37 °C, 3 h		

Figure 5.3: SDS-PAGE of the membrane fractions of the expression test of PgaA in the pColdIV vector. *E.coli* strains and expression temperature were varied. The induction was carried out with 1 mM IPTG. The best expression was achieved with the C41(DE3) strain at 37 °C after 3 h induction (lane 8). Sample preparation was done as described in section 2.2.1.

coli strains BL21(DE3), BL21(DE3)pLysS, Rosetta(DE3), C41(DE3), C43(DE3), ArcticExpress(DE3) and BL21(DE3)omp8/pLysS were tested at two different temperatures of 37 °C and 18 °C. The expression was induced with 1 mM IPTG. The analysis of the expression was performed by collecting the outer membrane fraction and yielded the best expression in the C41(DE3) strain at 37 °C after 3 h induction (Figure 5.3). Nevertheless, the expression using the pET28a vector was much better compared to the pColdIV vector.

N-terminal sequencing revealed that the signal sequence of PgaA was not cleaved and therefore the transport to the outer membrane might be disturbed. This might be a reason for the bad expression level of PgaA in the pET28a and pColdIV vector. To avoid this problem, the PgaA signal sequence was substituted by other signal sequences, which are known to express high amounts of membrane proteins.
First, the expression was tested in the pET22b/PgaA vector, where the *pelB* signal sequence was used. The different *E. coli* strains BL21(DE3), BL21(DE3)pLysS, Rosetta(DE3), C43(DE3), ArcticExpress(DE3) and BL21(DE3)omp8/pLysS were tested at two different temperatures of 37 °C and 18 °C. The expression was induced with 1 mM IPTG. The best expression was achieved in the BL21(DE3)omp8/pLysS strain at 37 °C after 3 h induction (Figure 5.4), but was not as much as obtained with the pET28a/PgaA vector.

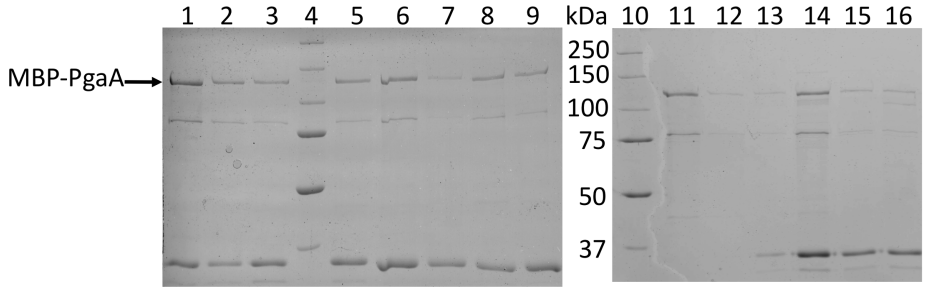


lane 1:	marker	lane 8:	C43(DE3), 37 °C, 3 h
lane 2:	BL21(DE3), 37 °C, 3 h	lane 9:	C43(DE3), 18 °C, 20 h
lane 3:	BL21(DE3), 18 °C, 20 h	lane 10:	ArcticExpress(DE3), 37 °C, 3 h
lane 4:	BL21(DE3)pLysS, 37 °C, 3 h	lane 11:	ArcticExpress(DE3), 18 °C, 20 h
lane 5:	BL21(DE3)pLysS, 18 °C, 20 h	lane 12:	BL21(DE3)omp8/pLysS, 37 °C, 3 h
lane 6:	Rosetta(DE3), 37 °C, 3 h	lane 13:	BL21(DE3)omp8/pLysS, 18 °C, 20 h
lane 7:	Rosetta(DE3), 18 °C, 20 h	lane 14:	marker

Figure 5.4: SDS-PAGE of the membrane fractions of the expression test of PgaA in the pET22b vector. *E.coli* strains and expression temperature were varied. The induction was carried out with 1 mM IPTG. The best expression was achieved in the BL21(DE3)omp8/pLysS strain at 37 °C after 3 h induction (lane 12). Sample preparation was done as described in section 2.2.1.

In a next step, the expression of PgaA was tested in the pMAL-p5X vector, where PgaA is fused to MBP and the signal sequence of MBP is used. In expression tests, the different *E. coli* strains BL21(DE3), BL21(DE3)pLysS, Rosetta(DE3), C41(DE3), C43(DE3), ArcticExpress(DE3) and BL21(DE3)omp8/pLysS were tested at two different temperatures of 37 °C and 18 °C. The expression was induced with 1 mM IPTG. The best expression of MBP-PgaA was obtained in BL21(DE3) cells at 37 °C after 3 h induction (Figure 5.5). However, the yield of the expression was not as good as with the pET28a vector.

In the expression vector pET40b/PgaA, PgaA is fused to the periplasmic enzyme DsbC that catalyzes the formation and isomerization of disulfide bonds. This system uses the signal sequence of DsbC. Expression tests were performed in the different *E. coli* strains BL21(DE3), BL21(DE3)pLysS, Rosetta(DE3), C43(DE3) and ArcticExpress(DE3) at two different temperatures of 37 °C and 18 °C. The expression was induced with 1 mM IPTG. The best results were achieved in BL21(DE3) and C43(DE3) cells at 18 °C and in Rosetta(DE3) cells at 37 °C (Figure 5.6). The expression was not as good as with the pET28a vector, but gave the best results with a different signal sequence than that of PgaA itself.
The expression was further refined by varying the IPTG concentration. Reduced IPTG concentrations of 100 or 10 μM had no influence in BL21(DE3) cells at 18 °C, but in C43(DE3)

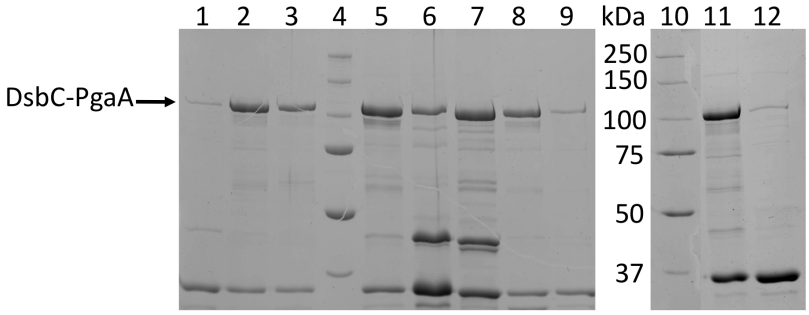


lane 1:	BL21(DE3), 37 °C, 3 h	lane 9:	Rosetta(DE3), 18 °C, 20 h
lane 2:	BL21(DE3)pLysS, 37 °C, 3 h	lane 10:	marker
lane 3:	C41(DE3), 37 °C, 3 h	lane 11:	BL21(DE3)omp8/pLysS, 37 °C, 3 h
lane 4:	marker	lane 12:	BL21(DE3)omp8/pLysS, 18 °C, 20 h
lane 5:	C43(DE3), 37 °C, 3 h	lane 13:	BL21(DE3)pLysS, 18 °C, 20 h
lane 6:	Rosetta(DE3), 37 °C, 3 h	lane 14:	C41(DE3), 18 °C, 20 h
lane 7:	Arctic(DE3), 37 °C, 3 h	lane 15:	C43(DE3), 18 °C, 20 h
lane 8:	BL21(DE3), 18 °C, 20 h	lane 16:	ArcticExpress(DE3), 18 °C, 20 h

Figure 5.5: SDS-PAGE of the membrane fractions of the expression test of MBP-PgaA in the pMal-p5X vector. *E.coli* strains and expression temperature were varied. The induction was carried out with 1 mM IPTG. The best expression was achieved in BL21(DE3) cells at 37 °C after 3 h induction (lane 1). Sample preparation was done as described in section 2.2.1.

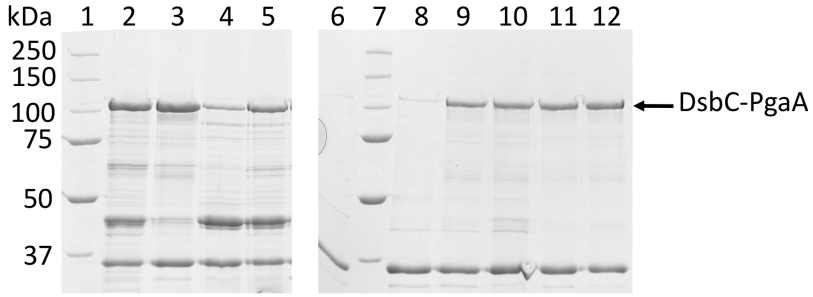
cells at 18 °C the expression was reduced compared to the induction with 1 mM IPTG (Figure 5.7). In Rosetta(DE3) cells at 37 °C using 10 μM IPTG for induction, PgaA was not expressed. With 100 μM IPTG, PgaA was expressed in same amounts at 37 °C and 30 °C, but the expression was lower than with 1 mM IPTG (Figure 5.7). The best expression level with the pET40b/PgaA was reached in BL21(DE3) cells at 18 °C with an IPTG concentration between 0.1-1 mM.

For large scale expression, two different vector systems were chosen, the pET28a/PgaA plasmid using the signal sequence of PgaA and the pET40b/PgaA plasmid using a the signal sequence of DsbC. Bacteria were cultivated on a 5 L scale with the conditions established in the expression tests. PgaA in the pET28a vector was expressed in BL21(DE3) cells at 37 °C. Induction was carried out with 1 mM IPTG for 3 h. PgaA in the pET40b vector was expressed in BL21(DE3) cells at 18 °C. Induction was carried out with $0.1 - 1$ mM IPTG for 20 h.



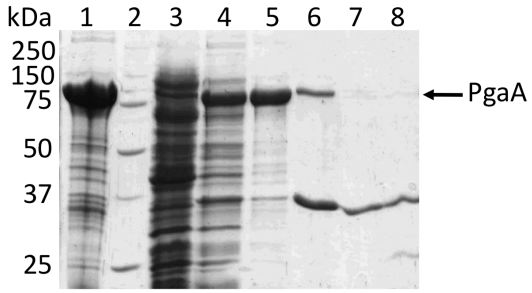
lane 1:	C43(DE3), 37 °C, 3 h	lane 7:	C43(DE3), 18 °C, 20 h
lane 2:	Rosetta(DE3), 37 °C, 3 h	lane 8:	Rosetta(DE3), 18 °C, 20 h
lane 3:	ArcticExpress(DE3), 37 °C, 3 h	lane 9:	ArcticExpress(DE3), 18 °C, 20 h
lane 4:	marker	lane 10:	marker
lane 5:	BL21(DE3), 18 °C, 20 h	lane 11:	BL21(DE3), 37 °C, 3 h
lane 6:	BL21(DE3)pLysS, 18 °C, 20 h	lane 12:	BL21(DE3)pLysS, 37 °C, 3 h

Figure 5.6: SDS-PAGE of the membrane fractions of the expression test of DsbC-PgaA in the pET40b vector. *E.coli* strains and expression temperature were varied. The induction was carried out with 1 mM IPTG. The best expression was achieved in BL21(DE3) and C43(DE3) cells at 18 °C and in Rosetta(DE3) cells at 37 °C (lane 5, 7, 2). Sample preparation was done as described in section 2.2.1.



lane 1:	marker	lane 7:	marker
lane 2:	BL21(DE3), 18 °C, 20 h, 10 μM IPTG	lane 8:	Rosetta(DE3), 37 °C, 5 h, 10 μM IPTG
lane 3:	BL21(DE3), 18 °C, 20 h, 100 μM IPTG	lane 9:	Rosetta(DE3), 37 °C, 3 h, 100 μM IPTG
lane 4:	C43(DE3), 18 °C, 20 h, 10 μM IPTG	lane 10:	Rosetta(DE3), 37 °C, 5 h, 100 μM IPTG
lane 5:	C43(DE3), 18 °C, 20 h, 100 μM IPTG	lane 11:	Rosetta(DE3), 30 °C, 3 h, 100 μM IPTG
lane 6:	Rosetta(DE3), 37 °C, 3 h, 10 μM IPTG	lane 12:	Rosetta(DE3), 30 °C, 5 h, 100 μM IPTG

Figure 5.7: SDS-PAGE of the refined expression test of DsbC-PgaA in the pET40b vector. The IPTG concentration was varied using the strains BL21(DE3) and C43(DE3) (lane 2 - 5). The expression in Rosetta(DE3) cells was refined by changing expression temperature and IPTG concentration (lane 6 -12). The best expression was achieved in BL21(DE3) cells at 18 °C with an IPTG concentration of 0.1-1 mM (lane 2 and 3). Sample preparation was done as described in section 2.2.1.



lane 1: inclusion bodies
lane 2: marker
lane 3: soluble fraction
lane 4: extraction with 1 % N-lauroyl sarcosine
lane 5: extraction with 0.05 % LDAO
lane 6: extraction with 0.1 % LDAO
lane 7 : extraction with 0.2 % LDAO
lane 8: extraction with 0.3 % LDAO

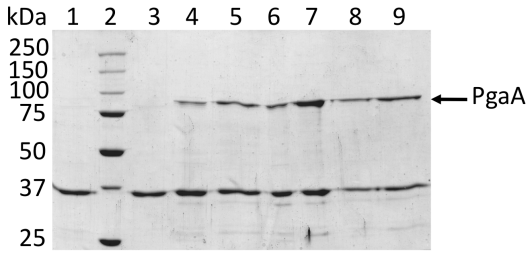
Figure 5.8: Extraction of PgaA from the outer membrane with the detergent LDAO. Inclusion bodies were resuspended in 1 % N-lauroyl sarcosine and loaded on the gel. 15 μL were loaded in every lane.

Purification of PgaA

The first step of the purification of PgaA was its extraction from the membrane. Therefore, the membrane fraction was first treated with N-lauroyl sarcosine to remove the inner membrane. In a second step, the proteins from the outer membrane were extracted with increasing detergent concentrations. In first extraction tests, different detergents were used (DM, LDAO, OPOE, OG, DDM, $C_{12}E_9$, Fos-12 and Cymal-6). With all detergents, PgaA could be extracted from the outer membrane in the first extraction step, which was probably caused by the large soluble domain of PgaA. The best result for PgaA extraction from the outer membrane could be achieved with LDAO (Figure 5.8). As it seen, on the SDS-PAGE of the extraction, PgaA was extracted from the inner membrane as well (Figure 5.8). This might be due to the saturation of the export machinery, which directs PgaA to the outer membrane.

Extraction was improved by varying the pH of the extraction buffer and the presence of salt. Therefore, the outer membrane fraction was resuspended in different extraction buffers (100 mM MES, pH 6.5, 100 mM Na_2HPO_4, pH 7.5, 100 mM Tris-HCl, pH 8.0 or 100 mM Bicine, pH 9.0 with and without 200 mM NaCl) and PgaA was extracted by adding 0.5 % LDAO. As it is observed in Figure 5.9, PgaA could not be extracted at pH 6.5. The best results were obtained at a pH of 8.0. The addition of salt always resulted in slightly higher extracted PgaA levels. Gel filtration experiments revealed that PgaA had the tendency to form aggregates without salt, so the extraction was performed at pH 8.0 in the presence of 200 mM NaCl using LDAO as detergent.

Extracted PgaA was further purified by using Ni-affinity chromatography with 100 mM Na_2HPO_4, pH 7.5, 200 mM NaCl, 0.1 % LDAO as buffer. A chromatogram of an exemplary run is shown in Figure 5.10 and the corresponding SDS-PAGE is shown in Figure 5.11. It can be seen that the binding of PgaA to the column is very weak, because it already elutes



lane 1: MES, pH 6.5
lane 2: marker
lane 3: MES, pH 6.5, NaCl
lane 4: Na_2HPO_4, pH 7.5
lane 5: Na_2HPO_4, pH 7.5, NaCl
lane 6: Tris-HCl, pH 8.0
lane 7 : Tris-HCl, pH 8.0, NaCl
lane 8: Bicine, pH 9.0
lane 9: Bicine, pH 9.0, NaCl

Figure 5.9: SDS-PAGE of the buffer screening for the extraction of PgaA from the outer membrane. PgaA was extracted with 0.5 % LDAO. The best result was obtained in Tris-HCl, pH 8.0 in the presence of salt (lane 7). 15 μL were loaded in every lane.

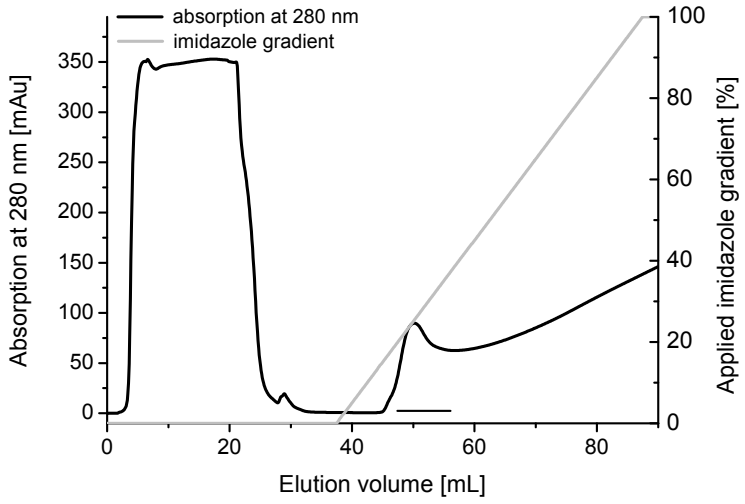
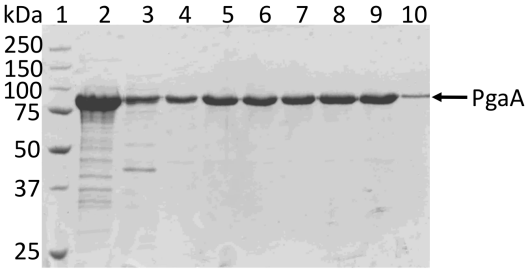
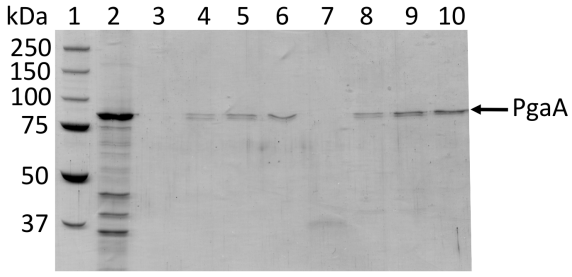


Figure 5.10: Chromatogram of PgaA purification by Ni-affinity chromatography using a 5 mL His-Trap column. PgaA elutes at an imidazole concentration of 50 mM. Fractions pooled and used for subsequent purification are marked by a line.



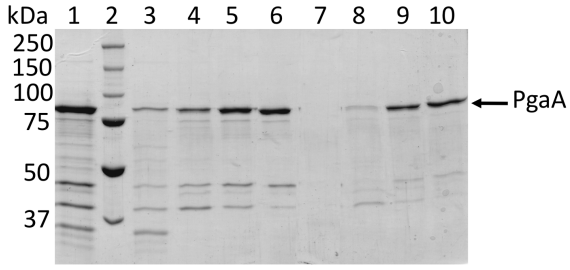
lane 1: marker
lane 2: flow through (0-28 mL)
lane 3: fraction 45.5-48 mL
lane 4-9: fractions 48-53 mL
lane 10: fraction 57-58.5 mL

Figure 5.11: SDS-PAGE of the PgaA purification by Ni-chromatography. 2 μL protein solution loaded were in lane 2 and 10 μL in lane 3-10.



100 mM Tris-HCl, pH 8.0, 200 mM NaCl

lane 1: marker
lane 2: load
lane 3-6: elution
lane 7: loading in circle for 18 h, flow through
lane 8-10: loading in circle for 18 h, elution



100 mM Na_2HPO_4, pH 8.0, 200 mM NaCl

lane 1: load
lane 2: marker
lane 3: flow through
lane 4-6: elution
lane 7: loading in circle for 18 h, flow through
lane 8-10: loading in circle for 18 h, elution

Figure 5.12: SDS-PAGE of the PgaA purification by Ni-chromatography using different buffers at pH 8.0 and different loading modes. With 100 mM Na_2HPO_4, pH 8.0, 200 mM NaCl and loading in circle for 18 h the best result was achieved. 10 μL protein solution were loaded in every lane.

at an imidazole concentration of 50 mM and most of the PgaA has not even bound to the column as it is found in the flow through.
To improve the binding of PgaA to the Ni column, different buffers were tested. Because best results for PgaA extraction from the membrane were achieved at pH 8.0, Tris-HCl and phosphate buffer at pH 8.0 in the presence of 200 mM NaCl and 0.1 % LDAO were used. This buffers were already used for the extraction step and yielded same amounts of extracted PgaA. The purification via the Ni-column was performed in two different ways. One run was done with the standard procedure and a second run was performed by loading PgaA sample in a circle for 18 h onto the column. The results are shown in Figure 5.12. The use of Tris buffer, decreased the binding of PgaA to the column drastically. In phosphate buffer, the binding of PgaA was much better and by loading in circle for 18 h all PgaA has bound to the column. Therefore, the purification of PgaA by Ni-chromatography was performed with 100 mM phosphate buffer, 200 mM NaCl, 0.1 % LDAO and loading to the Ni column was done in circle for 18 h to obtain the best result.

After Ni-chromatography, PgaA was subjected to gel filtration as a final purification step. PgaA ran as one single peak with a elution volume of 12.1 mL and formed no aggregates (Figure 5.13 and Figure 5.14). The final yield of the purification was about 0.1 mg/L culture.

The stability of the purified protein was tested by gel filtration experiments after storage for four weeks at 4 °C and storage at −20 °C with and without 20 % glycerol added. The gel

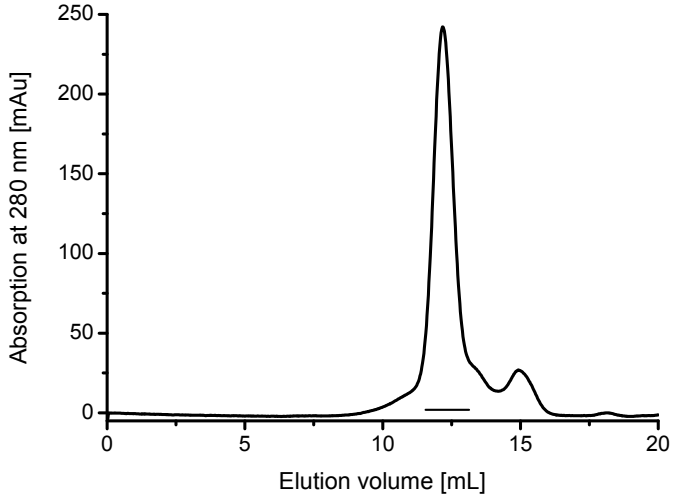


Figure 5.13: Chromatogram of PgaA purification by gel filtration chromatography using a Superdex 200 10/30 column. Pooled PgaA fractions are marked by a line.

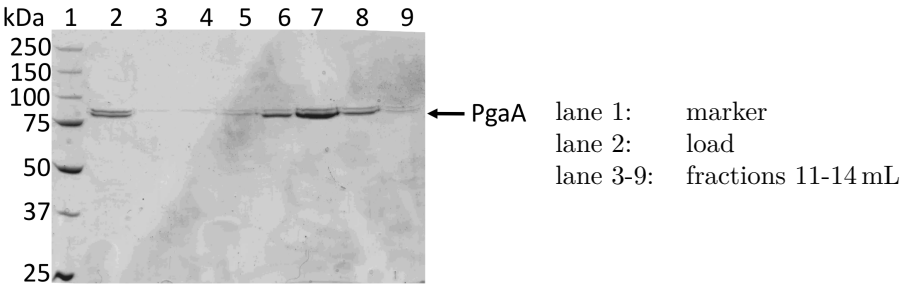


Figure 5.14: SDS-PAGE of the PgaA purification by gel filtration. 1 μL protein solution was loaded in lane 2 and 10 μL in lane 3-9.

filtration profile was identical for all tested samples (Figure 5.15). PgaA elutes as one single peak and did not form any aggregates. Thus, purified PgaA can be stored at least for four weeks at 4 ℃ and also the freezing and thawing of samples does not affect the stability of PgaA.

Purification of PgaA as DsbC fusion

The extraction of PgaA-DsbC was done like for PgaA using 100 mM Na_2HPO_4 pH 8.0, 200 mM NaCl as buffer and LDAO as detergent. The extraction behavior was the same as for PgaA, meaning that it was extracted in the first step (Figure 5.16). Afterwards, a Ni-chromatography was performed. A chromatogram is shown in Figure 5.17 and the corresponding SDS-PAGE is shown in Figure 5.18. PgaA is eluting at an imidazole concentration of 125 mM. After Ni-chromatography, PgaA-DsbC was subjected to gel filtration. As it can

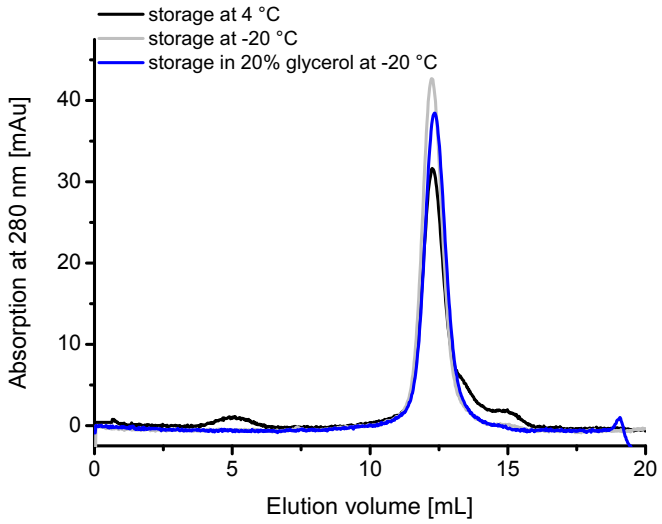


Figure 5.15: Chromatogram of PgaA samples stored for four weeks at 4 °C (black line) or at −20 °C with (blue line) and without 20 % glycerol added (grey line). PgaA elutes as one single peak under all storage conditions and has formed no aggregates.

be seen in the chromatogram (Figure 5.19) and on the corresponding SDS-PAGE (Figure 5.20), PgaA-DsbC does not elute homogeneously. The PgaA-DsbC elution is distributed over a range of 4.5 mL and elutes at least in two peaks. Because the purified PgaA-DsbC was not homogeneous, the work on this construct was stopped.

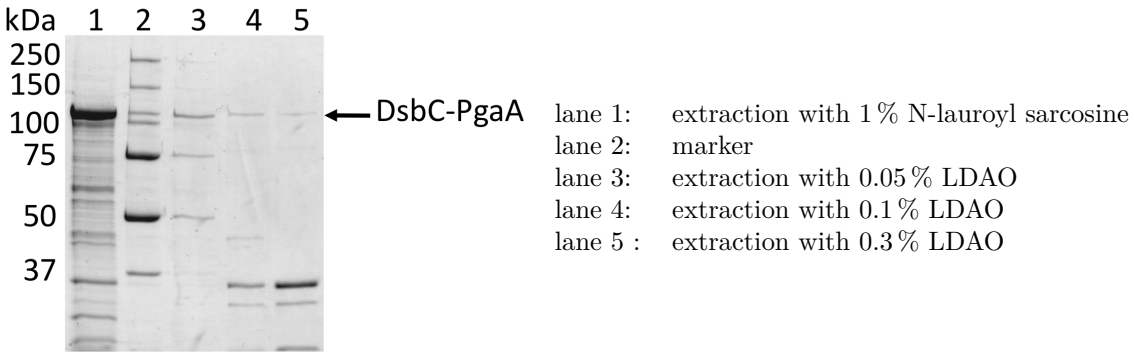


Figure 5.16: Extraction of DsbC-PgaA from the outer membrane with the detergent LDAO. 15 µL were loaded in every lane.

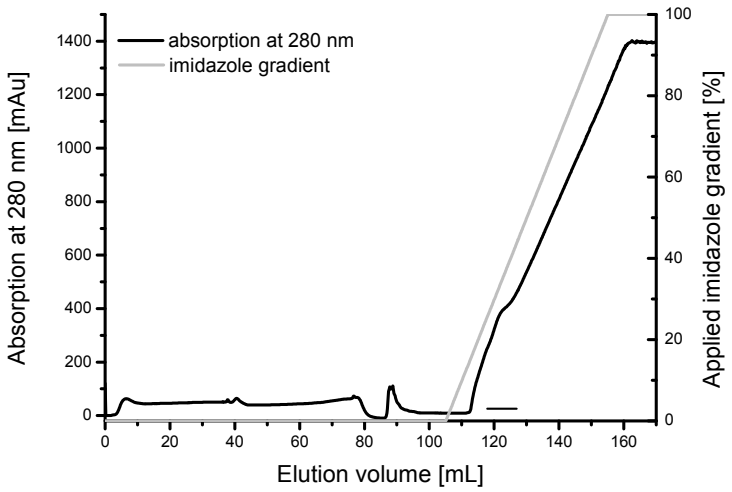
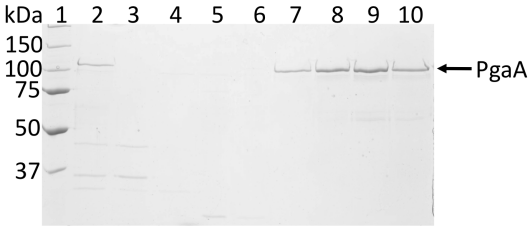


Figure 5.17: Chromatogram of PgaA-DsbC purification by Ni-affinity chromatography using a 5 mL HisTrap column. PgaA elutes at an imidazole concentration of 125 mM. Fractions pooled and used for subsequent purification are marked by a line.



lane 1:	marker
lane 2:	load
lane 3:	flow through (0-28 mL)
lane 4:	wash 85-105 mL
lane 5-6:	fractions 115.5-118.5 mL
lane 7-10:	fractions 118.5-127 mL

Figure 5.18: SDS-PAGE of the PgaA-DsbC purification by Ni-chromatography. 2 μL protein solution were loaded in lane 2-4 and 10 μL in lane 5-10.

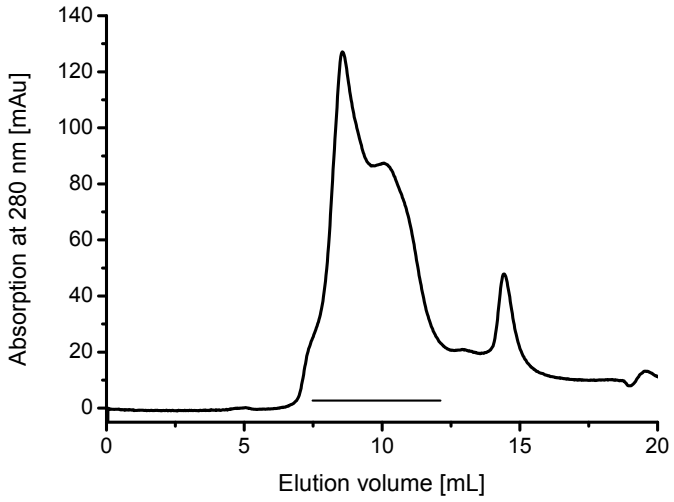


Figure 5.19: Chromatogram of PgaA-DsbC purification by gel filtration chromatography using a Superdex 200 10/30 column. PgaA-DsbC containing fractions are marked by a line.

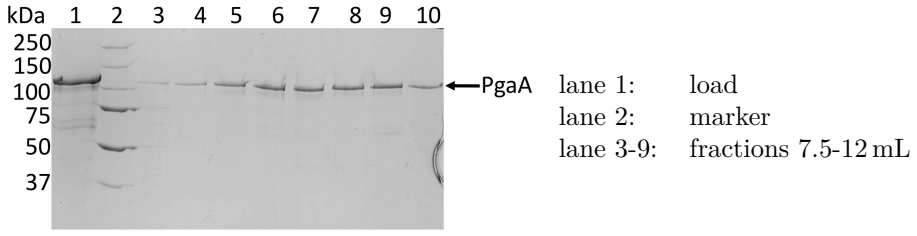
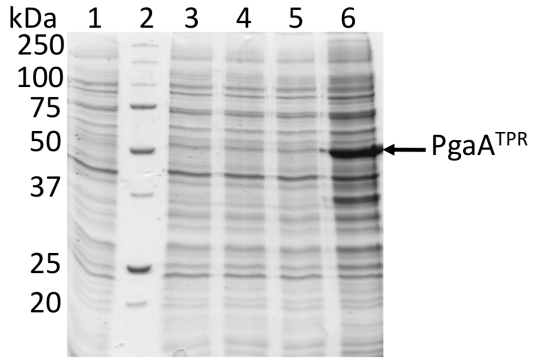


Figure 5.20: SDS-PAGE of the PgaA-DsbC purification by gel filtration. 1 μL protein solution was loaded in lane 1 and 10 μL in lane 3-9.



lane 1: before induction
lane 2: marker
lane 3: 37 °C, 2 h after induction
lane 4: 37 °C, 3 h after induction
lane 5: 37 °C, 4 h after induction
lane 6: 18 °C, 20 h after induction

Figure 5.21: SDS-PAGE of the soluble fractions of the expression test of $PgaA^{TPR}$ in the pET28a vector in Rosetta(DE3) cells. The expression temperature was 37 °C or 18 °C and the induction was carried out with 1 mM IPTG. The best expression was achieved at 18 °C after 20 h (lane 6).

Summary of the purification of PgaA

PgaA was purified by a stepwise extraction from the outer membrane in 100 mM Na_2HPO_4, pH 8.0, 200 mM NaCl buffer with the detergent LDAO. Afterwards, a two step-purification consisting of Ni-affinity and size exclusion chromatography was used. The Ni-column was performed in 100 mM Na_2HPO_4, pH 8.0, 200 mM NaCl, 0.1 % LDAO buffer by loading the sample to the Ni column in circle for 18 h. The final gel filtration yielded pure and homogenous protein, which could be stored for at least 4 weeks at 4 °C or −20 °C without forming aggregates. The yield of the purification was about 0.1 mg/L culture. The purification of PgaA as DsbC fusion protein resulted in heterogeneous protein samples and was therefore not continued.

Expression and Purification of $PgaA^{TPR}$

Expression tests for $PgaA^{TPR}$ were performed in the different *E. coli* strains BL21(DE3), BL21(DE3)pLysS, Rosetta(DE3), C43(DE3) and ArcticExpress(DE3) at two different temperatures of 37 °C and 18 °C. The expression was induced with 1 mM IPTG. The best soluble expression was achieved in the Rosetta(DE3) strain at 18 °C 20 h after induction. Exemplarily, an expression gel is shown in Figure 5.21. For protein production, bacteria was cultivated on a 5 L scale with the conditions established in the expression tests.

For purification, a two step-procedure consisting of Ni-affinity and size exclusion chromatography was used. After cell lysis with French press, Ni-affinity chromatography using a Histrap column was performed. A chromatogram of a typical run is shown in Figure 5.22 and the corresponding SDS-PAGE is shown in Figure 5.23. The protein elutes at an imidazole concentration of 165 mM. $PgaA^{TPR}$ containing fractions were pooled and subsequently

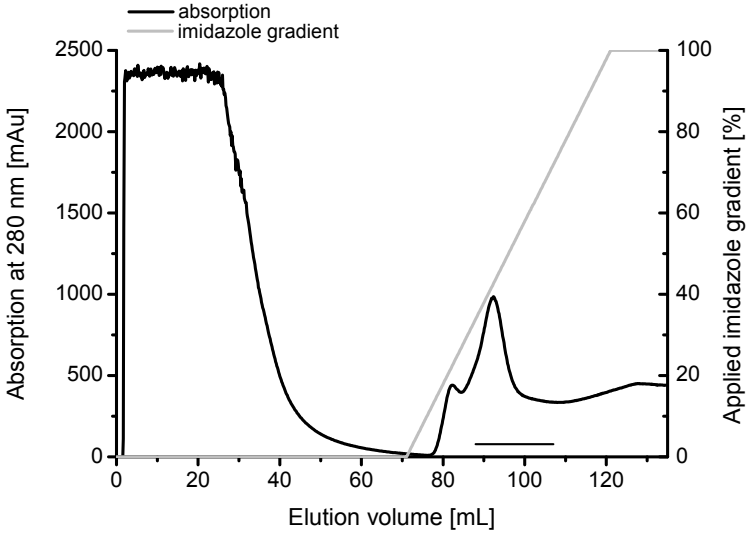
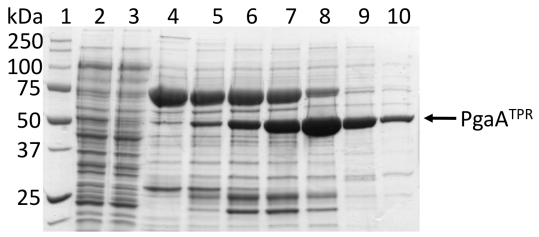


Figure 5.22: Chromatogram of PgaA^{TPR} purification by Ni-affinity chromatography using a 5 mL HisTrap column. PgaA^{TPR} elutes at an imidazole concentration of 165 mM. Fractions pooled and used for subsequent purification are marked by a line.



lane 1:	marker
lane 2:	load
lane 3:	flow through (0-28 mL)
lane 4:	fraction 82-83 mL
lane 5:	fraction 84-85 mL
lane 6:	fraction 87-88 mL
lane 7:	fraction 89-90 mL
lane 8:	fraction 92-93 mL
lane 9:	fraction 97-98 mL
lane 10:	fraction 106-107 mL

Figure 5.23: SDS-PAGE of the PgaA^{TPR} purification by Ni-chromatography. 2 μL protein solution were loaded in lane 2 and 3 and 10 μL in lane 4-10.

loaded on a gel filtration column. PgaA^{TPR} ran as one single peak with a elution volume of 195.5 mL (Figure 5.24 and 5.25). According to a calibration curve for the used column, this corresponds to a molecular mass of 61.3 kDa. Assuming the mass of one protein chain of 55.8 kDa, this clearly indicates that PgaA^{TPR} is monomeric. The final yield of the purification was about 3.5 mg/L culture.

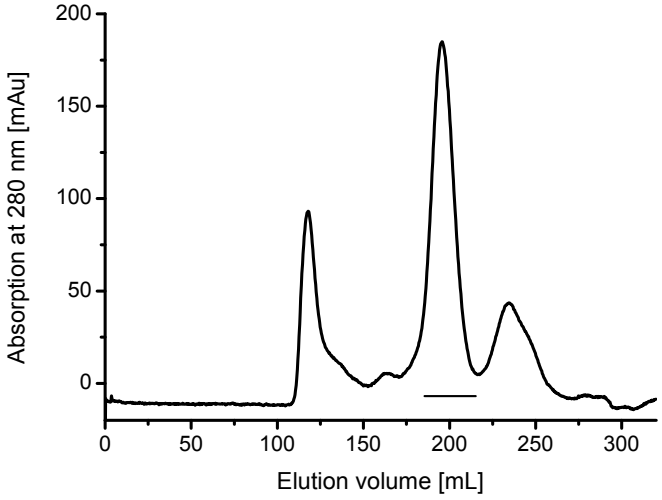
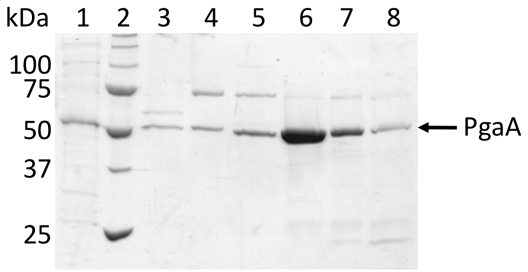


Figure 5.24: Chromatogram of PgaA^{TPR} purification by gel filtration chromatography using a HiLoad-26/60-Superdex-200 prep grade column. PgaA^{TPR} was separated from minor aggregates, which elute at about 120 mL. Pooled PgaA^{TPR} fractions are marked by a line.



lane 1: fraction 115-120 mL
lane 2: marker
lane 3: fraction 135-140 mL
lane 4: fraction 165-170 mL
lane 5-8: fractions 185-215 mL

Figure 5.25: SDS-PAGE of the PgaA^{TPR} purification by gel filtration. 10 μL protein solution was loaded in each lane.

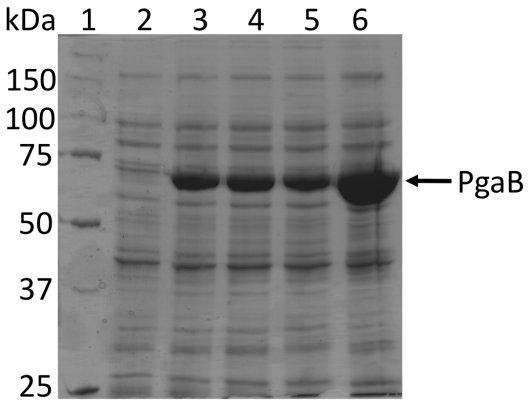


Figure 5.26: SDS-PAGE of the soluble fractions of the expression test of PgaB in the pET21b vector in BL21(DE3) cells. The expression temperature was 37 °C or 18 °C and the induction was carried out with 1 mM IPTG. The best expression is achieved at 18 °C 20 h after induction (lane 6).

5.1.3 Expression and Purification of PgaB

The constructs pET21b/PgaB and pET28a/PgaB in different *E. coli* strains BL21(DE3), BL21(DE3)pLysS, Rosetta(DE3), C43(DE3) and ArcticExpress(DE3) were tested for PgaB expression at two different temperatures of 37 °C and 18 °C. The expression was induced with 1 mM IPTG. In most of the tested conditions, soluble PgaB was highly expressed. The best expression was achieved for pET21b/PgaB in BL21(DE3) cells at 18 °C 20 h after induction and for pET28a/PgaB in the Rosetta(DE3) strain at 18 °C 16 h after induction. Exemplarily, expression gels are shown in Figure 5.26 and 5.27.
For protein production, bacteria was cultivated on a 200 mL scale with the conditions established in the expression tests. Depending on further experiment, PgaB was expressed with a N- or C-terminal Histag.

As the purification procedure is the same for both constructs, only the PgaB purification from the construct pET28a/PgaB is shown. The purification protocol for PgaB consists of a Ni-affinity and a size-exclusion chromatography. After cell lysis with French press, Ni-affinity chromatography using a Histrap column was performed. A chromatogram of a typical run is shown in Figure 5.28 and the corresponding SDS-PAGE is shown in Figure 5.29. The protein elutes at an imidazole concentration of 125 mM. After Ni-chromatography, PgaB was subjected to gel filtration as a final purification step. PgaB ran as one single peak with an elution volume of 204.5 mL (Figure 5.30 and 5.31). According to the calibration curve for the used column, this corresponds to a molecular mass of 46.3 kDa. Assuming the mass of one protein chain of 75.0 kDa, this indicates that PgaB is a monomeric protein. The fact that the determined mass of PgaB is much smaller than the theoretical mass, indicates

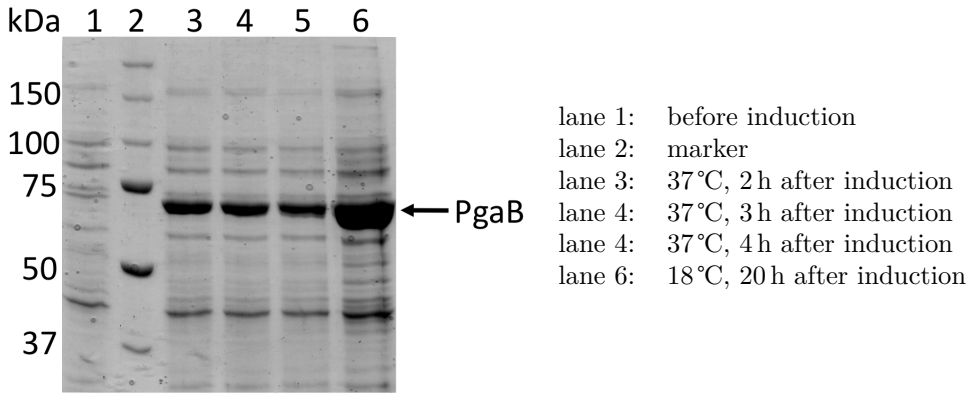


Figure 5.27: SDS-PAGE of the soluble fractions of the expression test of PgaB in the pET28a vector in Rosetta(DE3) cells. The expression temperature was 37 °C or 18 °C and the induction was carried out with 1 mM IPTG. The best expression is achieved at 18 °C 20 h after induction (lane 6).

towards a very compact protein. A degradation can be excluded, because on SDS-PAGE, PgaB runs at a size of about 75 kDa and the Histag of both constructs could be detected in a Western blot (data not shown). The final yield of the purification was about 250 mg/L culture using the pET28a/PgaB construct and 340 mg/L for the pET21b/PgaB construct.

The stability of the purified protein was tested by gel filtration experiments after storage for seven weeks at 4 °C and storage at −20 °C with and without 20 % glycerol added. Upon thawing the sample, that was stored at −20 °C without 20 % glycerol, PgaB was heavily precipitating. The two other samples were clear and were analyzed by gel filtration (Figure 5.32). In both cases, PgaB eluted as one single peak and has formed no aggregates. This means that purified PgaB can be stored at 4 °C for at least seven weeks and also the freezing and thawing of samples, when 20 % glycerol is added, does not affect the stability of PgaB.

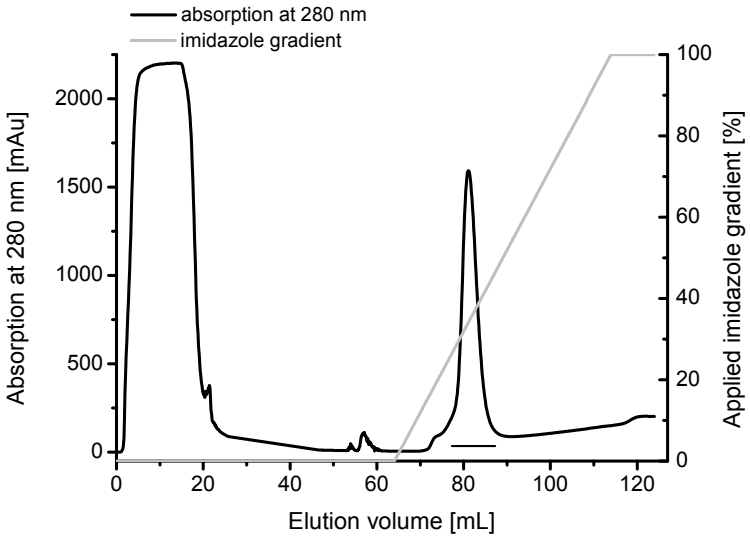
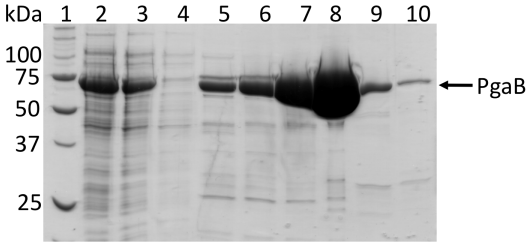


Figure 5.28: Chromatogram of PgaB purification by Ni-affinity chromatography using a 5 mL His-Trap column. PgaB elutes at an imidazole concentration of 125 mM. Fractions pooled and used for subsequent purification are marked by a line.



lane 1:	marker
lane 2:	cell extract
lane 3:	load
lane 4:	flow through (0-21 mL)
lane 5:	fraction 76-77 mL
lane 6-9:	fractions 77-87.5 mL
lane 10:	fraction 90-91 mL

Figure 5.29: SDS-PAGE of the PgaB purification by Ni-chromatography. 2 μL protein solution were loaded in lane 2-4 and 10 μL in lane 5-10.

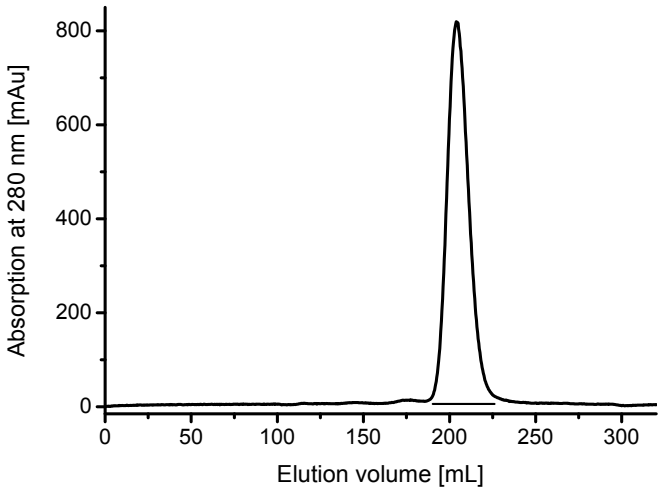
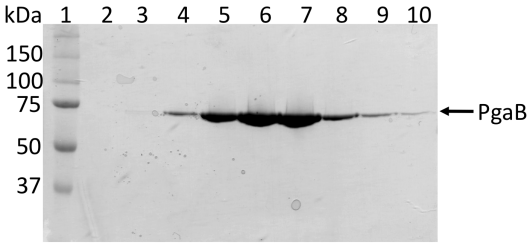


Figure 5.30: Chromatogram of PgaB purification by gel filtration chromatography using a HiLoad-26/60-Superdex-200 prep grade column. PgaB elutes as a single peak. Pooled PgaB fractions are marked by a line.



lane 1:	marker
lane 2:	fraction 110-115 mL
lane 3:	fraction 170-175 mL
lane 4-9:	fractions 190-225 mL
lane 10:	fraction 230-235 mL

Figure 5.31: SDS-PAGE of the PgaB purification by gel filtration. 10 μL protein solution were loaded in each lane.

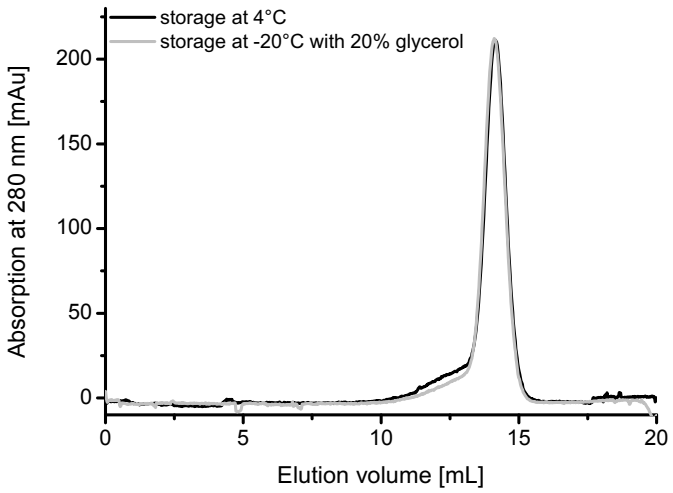


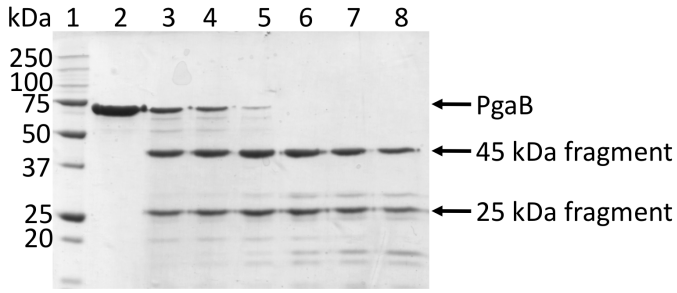
Figure 5.32: Chromatogram of PgaB samples stored for seven weeks at 4 °C or at −20 °C with 20 % glycerol added. PgaB elutes as one single peak under both storage conditions and has formed no aggregates.

5.1.4 Crystallization attempts of PgaA

As the yield of the purification of PgaA was very low, only a few attempts were made to get full-length PgaA crystals. The commercial crystallization screens Crystal Screen, Index and Mb-class were tested at a PgaA concentrations of 5.5 mg/mL. No crystals were obtained under these conditions. Because most of the drops stayed clear, the protein concentrations must be increased in further experiments.
For $PgaA^{TPR}$, the crystallization screens Crystal Screen and Index were used to setup crystallization experiments at a protein concentration of 10 mg/mL. None of the tested conditions yielded protein crystals. In nearly all drops $PgaA^{TPR}$ was heavily precipitated. As the protein solution of $PgaA^{TPR}$ itself, which was not used for crystallization, also starts to precipitate at 10 mg/mL, it is likely that $PgaA^{TPR}$ is not stable at such concentration.

5.1.5 Crystallization attempts of PgaB

A lot of attempts were made to obtain PgaB crystals. Commercial crystallization screens were tested at different protein concentrations in 96-well sitting drop set-ups. All trials are listed in Table 5.3. In the first attempts, protein concentration was too high, which resulted in precipitation in nearly every condition. In further experiments, the PgaB concentration was decreased. Furthermore, PgaB with a C-terminal or a N-terminal Histag were both used, because the Histag might influence the crystallization behavior.
Next, limited proteolysis was tested, which should reveal stable domains without unstruc-



lane 1:	marker
lane 2:	0 min
lane 3:	1 min
lane 4:	2 min
lane 5:	5 min
lane 6:	15 min
lane 7:	30 min
lane 8:	60 min

Figure 5.33: SDS-PAGE of the digestion of PgaB with trypsin. Samples after different time points were loaded. A 1:200 dilution of trypsin was used. PgaB was cut in two fragments of about 45 kDa and 25 kDa.

tured and flexible termini. Therefore, different proteases (pepsin, proteinaseK, trypsin) were tested as described in section 2.3.7. Pepsin was not able to cleave PgaB, maybe due to the high pH of 7.5 for pepsin. ProteinaseK degrades PgaB completely into a number of small pieces. By analyzing the SDS-PAGE of the proteolysis with trypsin, PgaB was cleaved in two fragments of about 45 kDa and 25 kDa (Figure 5.33). To produce larger amounts of cleaved PgaB the trypsin digestion was scaled up (section 5.3) and the two protein fragments were tried to separate by gel filtration without success. Because no separation could be achieved by using a Ni-affinity chromatography or an ion-exchange chromatography using a MonoQ column as well, the two parts must still stick together indicating that trypsin cleavage probably occurs in a loop and does not disturb the tertiary structure of PgaB. As this could affect the crystallization behavior, the trypsin treated PgaB was used in crystallization experiments, which are listed in Table 5.3. However, no crystals could be obtained under these conditions either.

Another method, which was used to facilitate crystallization of PgaB, was the methylation of surface lysines. Therefore, PgaB was methylated as described in section 2.3.6 and purified from aggregates by gel filtration. Although the conditions for the methylation were rather harsh, PgaB did not form aggregates (Figure 5.34) and the yield was with 73 % quite high. The fractions marked with a line in the chromatogram were used for crystallization experiments. However, no crystals were obtained with the methylated form of PgaB either.

Table 5.3: Crystallization trails for PgaB.

PgaB	Histag	Concentration [mg/mL]	Used screens
wildtype	N-terminal	20, 80, 200	Crystal Screen, Classics light Suite, PEGs Suite, $AmSO_4$ Suite, Wizard I+II, Index, SaltRx
wildtype	N-terminal	5, 10, 50	PACT premier, JCSG-*plus*, Clear Strategy Screen I, Clear Strategy Screen II
wildtype	C-terminal	10, 20, 50	PACT premier, JCSG-*plus*, Structure screen, PEGs Suite, Index
Trypsin digested	N-terminal	10, 20	PACT premier, JCSG-*plus*, Structure screen, Index, MbClass
Methylated	C-terminal	10, 20, 40	PACT premier, JCSG-*plus*, Structure screen, Index

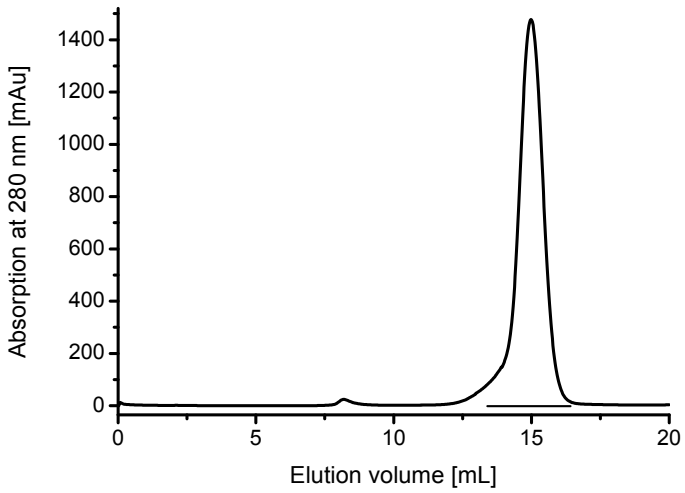


Figure 5.34: Chromatogram of PgaB after methylation of surface lysines. The Superdex-200 10/30 column was used. PgaB has not formed aggregates during the methylation reaction. The pooled fractions, which were used for crystallization experiments, are marked by a black line.

5.1.6 Investigations on PgaA-PgaB complex formation

As PgaB as well as PgaA are needed for the export of the poly-1,6-GlcNAc chain, it is likely that PgaB interacts with the TPR domain of PgaA, which is known to be involved in protein-protein interactions in other TPR-domain containing proteins. To prove this interaction, gel filtration experiments were performed as described in section 2.3.5. Therefore, $PgaA^{TPR}$ and PgaB were incubated in a molar ratio of 1:1 and 1:3 in 20 mM Tris-HCl, pH 8.0 150 mM NaCl and analyzed by gel filtration. In further experiments, 1 mM $MgCl_2$ or 1 mM GlcNAc was added for incubation to the protein mixture. In the chromatogram of all tested samples, no complex formation could be seen (Figure 5.35). PgaA and PgaB elute at the same position, as if they were loaded separately.
The fact that no complex formation could be detected with gel filtration experiments, might be due to the instability of the protein complex during the experiment, especially because it is diluted. Another reason could be, that an additional interaction partner is missing. This might be the poly-1,6-GlcNAc chain that can induce conformational changes after binding to PgaA or PgaB allowing complex formation. The experiment should be repeated in the presence of poly-1,6-GlcNAc. In addition, other methods like ITC or SPR should be used to analyze a possible PgaA-PgaB interaction.

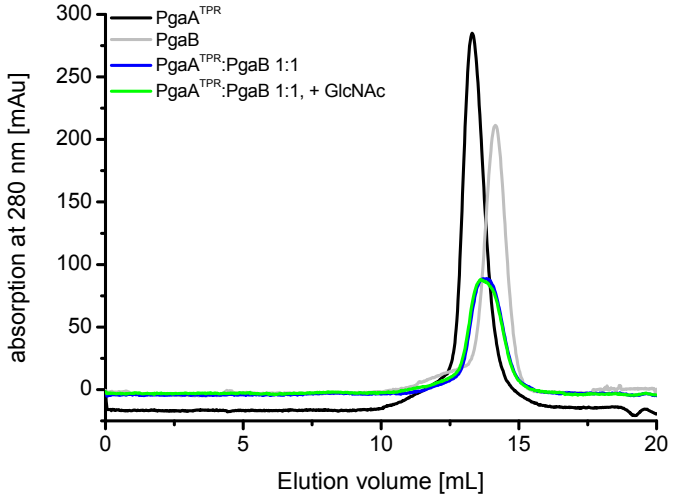


Figure 5.35: Gel filtration chromatograms of the investigation of complex formation of $PgaA^{TPR}$ and PgaB. $PgaA^{TPR}$ (black) and PgaB (grey) alone as well as samples containing both proteins in the presence (green) and absence of GlcNAc (blue) are shown. No complex formation could be seen, because in the run of the complex samples no additional peak at a smaller elution volume occur. The Superdex-200 10/30 column was used.

5.1.7 Activity of PgaB

PgaB shows sequence homology to other deacetylases and sequence alignments showed that all residues, which are involved in the deacetylation reaction, are present in PgaB. Therefore, experiments were carried out to prove that PgaB has deacetylation activity *in vitro*. For the realization of an enzymatic assay, the substrate of PgaB, poly-1,6-GlcNAc is needed. As poly-1,6-GlcNAc is not commercially available, it was enzymatically produced by PgaC and PgaD with a protocol developed by S. Steiner (see section 2.3.4). In this procedure, membrane fragments containing PgaC and PgaD were incubated with UDP-GlcNAc in the presence of c-di-GMP to produce poly-1,6-GlcNAc. Afterwards, the membranes can be removed by centrifugation and the produced poly-1,6-GlcNAc can be used as a substrate for PgaB in subsequent experiments. The synthesis of poly-1,6-GlcNAc by PgaC and PgaD can also be done in a one-step reaction together with PgaB. This approach was used, because it has several advantages. The fact that the chain length, which is required for PgaB activity is unknown, is not important, if PgaB is present during the poly-1,6-GlcNAc synthesis, because it is faced with all different chain lengths. The problem that poly-1,6-GlcNAc with more than ten monomers is not soluble, is also avoided. PgaB can perform the reaction and if afterwards the chain length will increase and the sugar will precipitate, the acetate remains soluble and can still be detected.
There are two possibilities to measure the activity of a deacetylases. One is the detection of

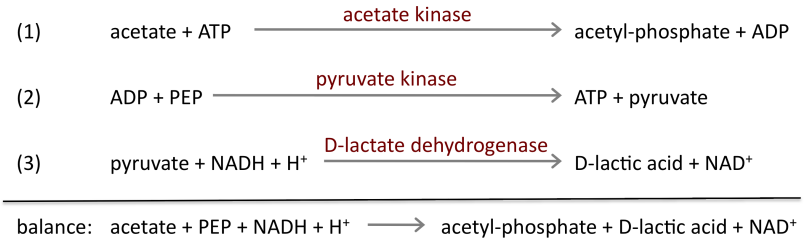


Figure 5.36: Principle of the acetate assay used in this thesis. (1) Acetate kinase in the presence of ATP converts acetate into acetyl-phosphate and ADP. (2) ADP formed in (1) is reconverted into ATP and pyruvate, by pyruvate kinase in the presence of phosphoenolpyruvate (PEP). (3) In the presence of the enzyme D-lactate dehydrogenase pyruvate is reduced to D-lactate by NADH with production of NAD^+. The amount of formed NAD^+ is stoichiometric with the amount of acetate. The NADH consumption is measured by the decrease in absorbance at 340 nm.

the generated free amines and the other is the measurement of the released acetate. In the last years, a fluorescence based assay was used to detect the free amines [155]. For PgaB, this assay is not suited, because if PgaB has deacetylated the sugar it precipitates, while the sugar chain is further growing by the action of PgaC, which means the free amines are not accessible for the detection. Therefore, an assay was used, which measures the released acetate. This can be done with a coupled assay, based on acetate-coenzyme A synthetase [156]. This method uses an indicator reaction catalyzed by L-malate dehydrogenase, which is in permanent equilibrium. Therefore, a non-stoichiometric correlation is observed to the acetate present in the sample. To overcome this problem, an alternative acetate detection assay was used in this thesis, based on the enzyme acetate kinase (Figure 5.36). As readout, the oxidation of NADH to NAD^+ was measured.

For activity measurements, PgaB was incubated for 16 h with membranes containing PgaC and PgaD together with UDP-GlcNAc in 50 mM HEPES, 50 mM NaCl. In a second sample, 5 mM $MgCl_2$ was added. Afterwards the proteins were precipitated by heating and removed by centrifugation. The cleared supernatants were analyzed with the acetate assay as described in section 2.3.4. PgaB activity could be measured independent of the presence of $MgCl_2$. The amount of acetate was such high, that the detection limit of the assay was reached.
In further experiments, the PgaB concentration and the reaction time will have to be decreased in order to measure the initial velocity of PgaB. Furthermore, the influence of different metal ions will have to be investigated, because it was shown that most of the deacetylases are metal dependent. For analysis of the chain length, which is needed for PgaB activity, poly-1,6-GlcNAc has to be chemically synthesized with different defined chain lengths.

5.2 Summary and outlook

For PgaA, an expression and purification protocol was established and resulted in stable and homogenous protein, that can be stored several weeks at 4 ℃ and can even be frozen. Initial crystallization trials did not yield protein crystals. However, further crystallization experiments with higher protein concentrations and in the presence of poly-1,6-GlcNAc should be done to provide insights in the sugar transport.
The soluble TPR domain of PgaA could be expressed and purified to homogeneity. Crystallization experiments did not yield protein crystals but should be repeated adding poly-1,6-GlcNAc of defined chain length and in the presence of PgaB to facilitate crystallization.

PgaB could be expressed in the cytosol by removing the lipid anchor and purified in very high amounts. The predicted deacetylase activity could be demonstrated *in vitro* with a coupled activity assay by detecting the released acetate. The substrate for PgaB was produced enzymatically by PgaC and PgaD in a one-step reaction in the presence of PgaB. In further experiments, the initial velocity will have to be measured. In addition, the metal dependency and the preferred chain length of PgaB needs to be investigated further. The developed activity assay is practicable in small volume in 96-well plates, so it is suitable for rapid screening of different reaction conditions and for the search of inhibitors for PgaB and PgaC.

To test the model that PgaA and PgaB are interacting with each other in order to allow the export of the poly-1,6-GlcNAc chain, it was tried to investigate the interaction by gel filtration experiments. An interaction could not be demonstrated, probably due to the limitations of the method. Another reason could be, that a poly-1,6-GlcNAc chain is required for the interaction of these two proteins. Complex formation should be further investigated in the presence of poly-1,6-GlcNAc and by using other techniques like ITC or SPR.

6 Appendix

6.1 Structure determination of KPC-2 in complex with a diazabicyclooctane inhibitor

6.1.1 KPC-2, a β-lactamase from *Klebsiella pneumoniae*

β-Lactamases catalyze the irreversible hydrolysis of the amide bond of β-lactams [157], and thereby protect the host organism against the lethal action of these antibiotics [158]. β-Lactamases are grouped into four molecular classes A, B, C and, D [159–161]. Members of class B are metalloproteins, whose activity is based on the presence of one or two zinc ions in their active site [162]. The majority of β-lactamases belong to the classes A, C and D, which use an active site serine for performing a two-step catalysis [163, 164]. β-Lactams acylate the active site serine forming a covalent acyl-enzyme intermediate. Subsequently, the acyl-enzyme intermediate is hydrolyzed by a water molecule. A schematic drawing of the reaction catalyzed is shown in Figure 6.1. Class A β-lactamases are monomeric enzymes, which consist of two domains, an all-α and an α/β domain [165–169]. The active site is located in the groove between these two domains.

KPC-2 is a class A β-lactamase that was originally identified in *K. pneumoniae*. KPC-2 is plasmid-encoded, is easily disseminated and is documented in numerous pathogens [170]. KPC-2 shows a wide substrate profile [171].

β-Lactams are the most common antibiotics in clinical use for the treatment of various infections caused by Gram-negative and Gram-positive pathogens [172–174]. The global problem

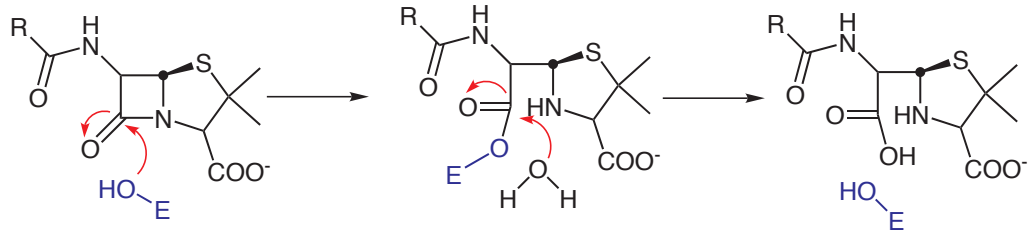


Figure 6.1: General catalytic pathway of active site serine β-lactamases. The first reaction step involves the nucleophilic attack of the active site serine to open the lactam ring. In a second step the deacylation by water hydrolyses the acyl-enzyme intermediate.

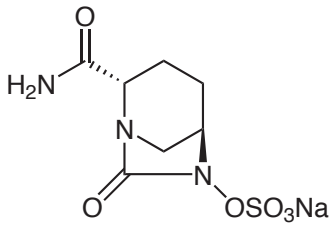


Figure 6.2: Chemical structure of the diazabicyclooctane inhibitor DBO-1.

of resistance to β-lactam antibiotics is growing into a serious problem for the therapy of infectious diseases [175, 176]. Therefore, β-lactam antibiotics have been administered with β-lactamase inhibitors [177]. The inhibitors of class A β-lactamases acylate the active site serine and form an acyl-enzyme intermediate which is refractory to hydrolysis. A new class of β-lactamase inhibitors are diazabicyclooctanes (DBOs) (Figure 6.2. The five-membered ring of a DBO contains an urea group that targets the active-site serine of the β-lactamase via a carbamylation reaction. This carbamate group is much more stable towards hydrolysis than the ester group generated in the reaction of classical β-lactam-derived β-lactamase inhibitors [178].
The structure of KPC-2 in complex with the diazabicyclooctane inhibitor DBO-1 was solved in order to elucidate the binding mode of this novel type of inhibitor and thereby to contribute to the optimization of this inhibitor class.

6.1.2 Crystallization

KPC-2 samples were received from Basilea Pharmaceutica and were further purified by size exclusion chromatography to remove potential aggregates and to exchange the buffer. The gel filtration chromatogram and the corresponding SDS gel are shown in Figure 6.3 and 6.4, respectively. Fractions containing KPC-2 were pooled and concentrated to 15-18 mg/mL.
The vapor-diffusion method was used for protein crystallization. First, the crystallization conditions described in ref. [169] were tested (Table 6.1). This was carried out in 24-well plates using hanging drops on siliconized cover slides. The drops were prepared by mixing 1 μL KPC-2 (17 mg/mL, in 20 mM Tris-HCl, pH 7.6) with 0.5 μL reservoir solution and were equilibrated against 500 μL reservoir buffer. Most conditions yielded crystals (Figure 6.5). Unfortunately, the crystallization drops contained many nuclei from which long, thin rods of crystals up to a size of 1000 x 40 x 40 μm^3 emanated. To reduce the amount of crystallization nuclei, precipitation concentration, pH value, protein concentration and crystallization temperature were varied (Table 6.2). Decreasing temperature, PEG 6000 or protein concentration, did not result in bigger crystals. The crystals grew slower, but there were still a lot of nuclei and they never reached the size reported in literature [169]. Single crystals

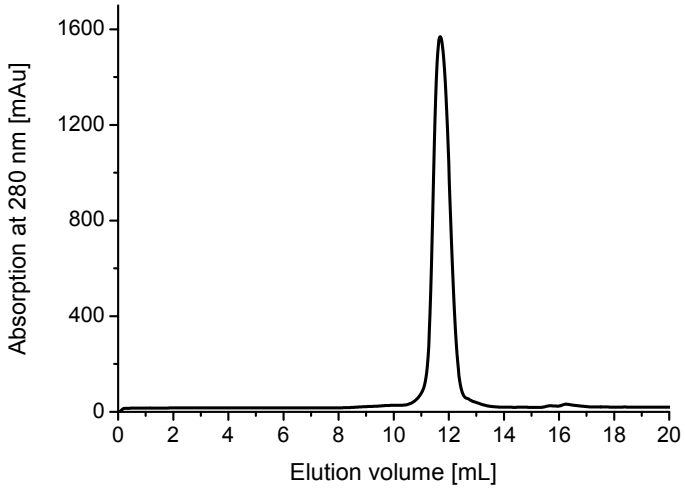
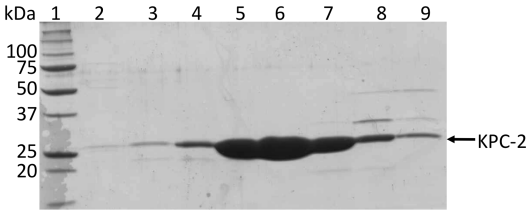


Figure 6.3: Chromatogram of the KPC-2 purification by gel filtration chromatography using a Superdex-75 10/30 column and 20 mM Tris-HCl, pH 7.6 as running buffer. KPC-2 elutes as a single peak with a elution volume of 11.7 mL.



lane 1:	marker
lane 2:	fraction 9-9.5 mL
lane 3:	fraction 9.5-10 mL
lane 4:	fraction 10-10.5 mL
lane 5:	fraction 10.5-11 mL
lane 6:	fraction 11-11.5 mL
lane 7:	fraction 11.5-12 mL
lane 8:	fraction 12-12.5 mL
lane 9:	fraction 12.5-13 mL
lane 10:	fraction 13-13.5 mL

Figure 6.4: SDS-PAGE of the KPC-2 purification by gel filtration. 5 μL protein solution were loaded in each lane.

of sufficient size were obtained by adding 10 mM spermine to the original crystallization condition. They grew to a final size of 400 x 40 x 40 μm^3 within 5 days (Figure 6.6).

In order to get KPC-2 inhibitor complex crystals, native KPC-2 crystals were soaked in a 5 μL soaking solution, containing 100 mM Bicine, pH 9.0, 16 % PEG 6000, 10 mM spermine and 2 mM inhibitor for 4 h. For cryoprotection, the soaked crystals were successively placed in soaking solution supplemented with 5, 10, 15 % glycerol with soaking times of 20 sec per step. Subsequently crystals were flash-frozen in liquid nitrogen.

Table 6.1: Original KPC-2 crystallization conditions from literature [169].

Parameter	Value
Protein concentration[a]	9 mg/mL
Reservoir solution	16 % PEG 6000, 100 mM Bicine, pH 9.0
Temperature	23 °C

[a] Final concentration in the crystallization drop.

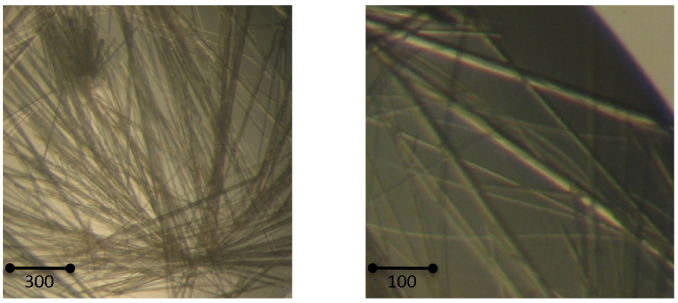


Figure 6.5: KPC-2 crystals obtained with the protocol described in ref [169]. Dimension are in μm.

Table 6.2: Parameter range explored to improve KPC-2 crystals.

Parameter	Value
Protein concentration[a]	6-13 mg/mL
Reservoir solution	4-22 % PEG 6000, 100 mM Bicine, pH 8.4-9.4
Temperature	23 °C and 4 °C

[a] Final concentration in the crystallization drop.

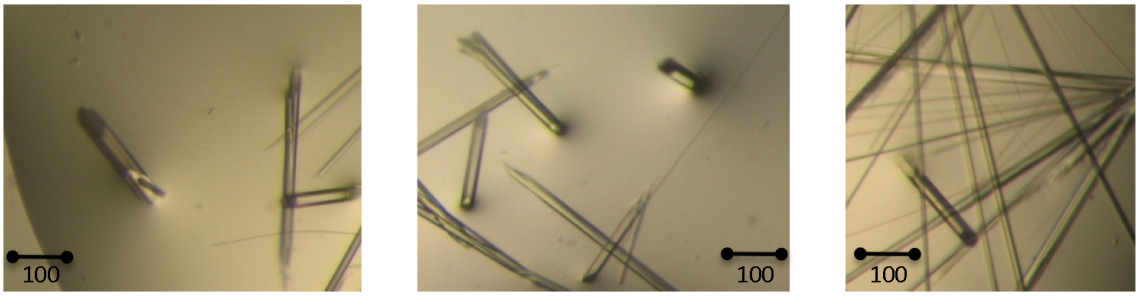


Figure 6.6: KPC-2 crystals obtained in the refined condition 100 mM Bicine, pH 9.0, 18 % PEG 6000, 10 mM spermine. Dimension are in μm.

6.1.3 Data collection and processing

Data collection was performed at the beamline PXI X06SA at the Swiss Light Source (Villigen, Switzerland). A complete dataset was collected to a resolution of 2.2 Å. For processing and scaling, the programs MOSFLM [109] and SCALA [110] were used. The parameters of the data collection and the statistics of the collected data set are summarized in Table 6.3. The crystal belongs to space group $P3_1$ with three molecules in the asymmetric unit and a solvent content of 50.5 %.

Table 6.3: Data collection parameters and statistics.

Wavelength [Å]	1.000
Detector	Pilatus 6M
Detector distance [mm]	300
Space group	$P3_1$
Cell axes [Å]	$a = b = 116.29, c = 52.79$
Angles [°]	$\alpha = \beta = 90, \gamma = 120$
Resolution [Å]	50-2.2 (2.1-2.2)
Observed reflections	95396
Unique reflections	40113
Multiplicity	2.4
R_{merge} [%] [a]	7.1 (30.2)
$I/\sigma(I)^a$	15.7 (2.4)
Completeness [%] [a]	100 (100)
Number of protomers per ASU	3
V_M [$Å^3 \cdot Da^{-1}$]	2.49
Solvent content [%]	50.5

[a] Values in parentheses refer to the highest resolution bin.

Unfortunately, the data set turned out to be twinned as indicated by the Merohedral Crystal Twinning Server [179]. The twinning operation was (h, -h-k, -l) and the twinning fraction 0.41. To continue structure elucidation, the dataset was de-twinned using DETWIN [124]. The KPC-2-inhibitor complex structure was determined with the de-twinned data by molecular replacement using MOLREP [180] with chain A of the KPC-2 structure (PDB entry 2OV5) as initial search model. MOLREP found a unique solution with three molecules in the asymmetric unit. The model was refined with REFMAC [114]. NCS restrains were applied between the three chains, and TLS refinement was carried out using each chain as an individual group. The model of the inhibitor was generated with PRODRG [181] and modeled into electron density with Coot [113]. Water molecules were added with ARP [182] and manually checked with Coot. The quality of the model was analyzed using PROCHECK [183].

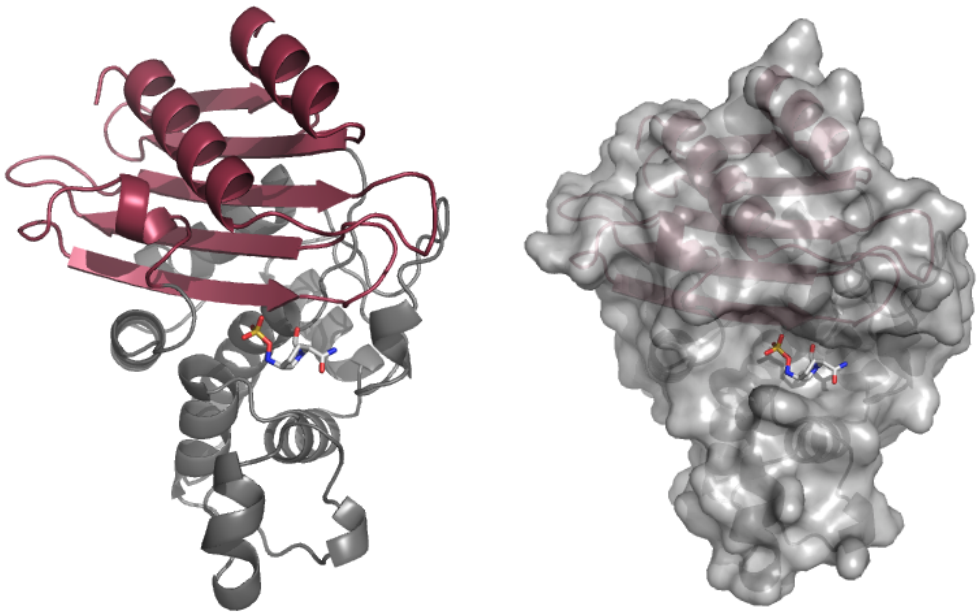
Table 6.4: Refinement statistics.

Resolution [Å]	2.2
Space group	$P3_1$
R [%]	23.6
R_{free} [%]	27.5
Number of atoms	
protein	11652
water molecules	93
inhibitor	51
B factors [Å^2]	
protein	40.4
inhibitor	44.2
water molecules	33.6
RMS deviations	
bond lengths [Å]	0.010
bond angles [°]	1.174
Ramachandran statistics [%]	
in favored regions	94.3
in allowed regions	5.1

6.1.4 Structure of the KPC-2-inhibitor complex

Considering that the data had to be de-twinned, the model has good quality indicators (Table 6.4) and 99.5 % of the residues had Φ and Ψ angles in the allowed regions of the Ramachandran plot. The final model is comprised of all 261 residues in all three chains, 93 water molecules and three inhibitor molecules. The complexed protein structure is virtually identical to the native structure with a RMS deviation of the Cα position of 0.06 Å. The three molecules in the asymmetric unit of the complex structure were considered identical, because the RMS Cα differences of the pairwise comparisons were below 0.06 Å. In the following, the structure of molecule A is described.

Clear density for the inhibitor was already visible in the first electron density map. The inhibitor appears covalently linked by an carbamate group to Ser70 and extends into the active site cleft (Figure 6.7). KPC-2 molecule A showed the best density. The omit electron density is shown in Figure 6.8. The average B factor for the inhibitor in KPC-2 molecule A is 36.2 Å^2, for molecules B and C the average B factors are somewhat higher (43.6 Å^2 and 52.8 Å^2). The inhibitor is stabilized via H-bonds with residues Ser70, Ser130, Asn132, Thr235 and Thr237 and hydrophobic contact with the residue Trp105 (Figure 6.8). The distances are depicted in Table 6.5. In the active site, a water molecule is found in a similar position to the deacylation water described in [169]. It is interacting with Glu166 and Asn170 and has a B factor of 37.4 Å^2. It is in a position to potentially hydrolyze the enzyme-inhibitor complex (Figure 6.9 and Table 6.6).



(A) (B)

Figure 6.7: Structure of the KPC-2 β-lactamase in complex with the inhibitor. (A) Ribbon diagram of KPC-2. One domain is shown in red and the other in grey. The inhibitor is covalently bound to Ser70 (shown in sticks). (B) Surface representation of KPC-2; the inhibitor is located in the catalytic cleft between both domains.

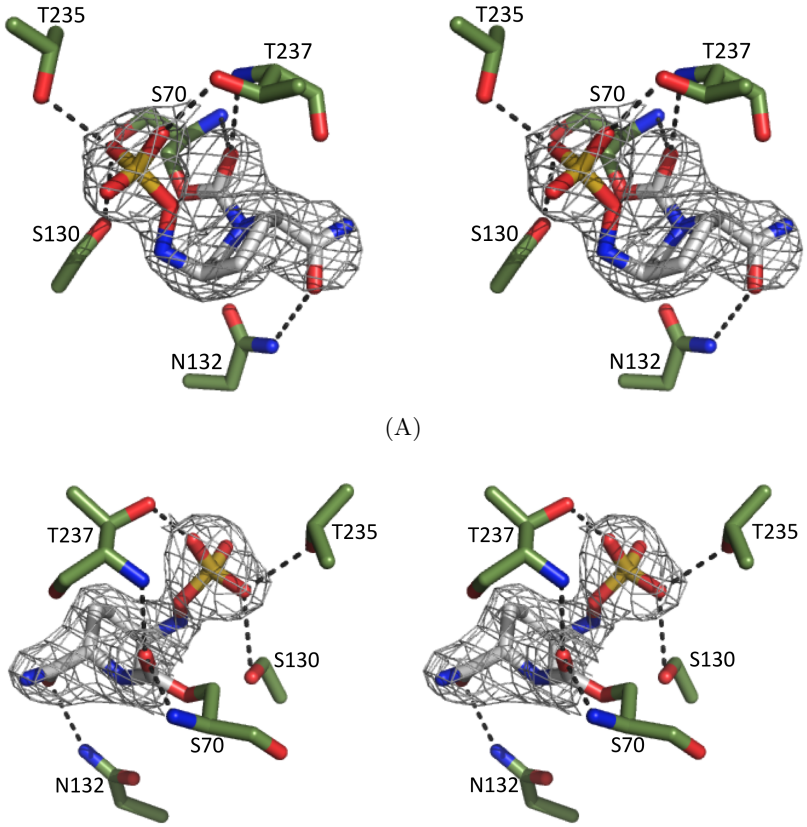


Figure 6.8: Stereo view of the active site of the KPC-2 β-lactamase with bound inhibitor. The interaction residues and the inhibitor are shown as sticks and the H-bonds are shown as dashed lines. The omit map for the ligand is countered at $2\,\sigma$. (A) View from the front like in Figure 6.7. (B) View from the back.

Table 6.5: Distances of the active site residues to the inhibitor.

Protein residue	Inhibitor	Distance [Å]
Ser170-N	O7	2.87
Ser130-O_γ	O3	3.32
Asn132-$N_{\delta 2}$	O14	3.21
Thr235-$O_{\gamma 1}$	O3	2.77
Thr237-$O_{\gamma 1}$	O4	2.42
Thr237-N	O7	2.95

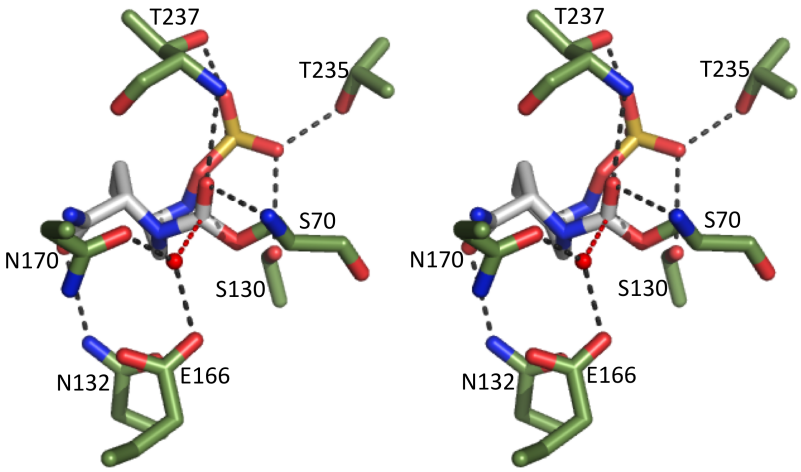


Figure 6.9: Stereo view of the active site of the KPC-2 β-lactamase with bound inhibitor and the de-acylation water (red small sphere). The interaction residues and the inhibitor are shown in full. The H-bonds are shown as black dashed lines. The nucleophilic attack of the water molecule is depicted as a red dashed line.

Table 6.6: Distances of the deacylation water to the active side residues and the inhibitor.

Interaction partner	Deacylation water	Distance [Å]
Glu166-O_{ϵ}	35	2.48
Asn170-$O_{\delta 1}$	35	2.32
Inhibitor-C6	35	2.89

Bibliography

[1] J. W. Costerton, L. Montanaro, and C. R. Arciola. Biofilm in implant infections: its production and regulation. *Int J Artif Organs* 28 (2005), pp. 1062–1068.

[2] R. Tamayo, J. T. Pratt, and A. Camilli. Roles of cyclic diguanylate in the regulation of bacterial pathogenesis. *Annu. Rev. Microbiol.* 61 (2007), pp. 131–148.

[3] P. Ross, H. Weinhouse, Y. Aloni, D. Michaeli, P. Weinberger-Ohana, R. Mayer, S. Braun, E. de Vroom, G. A. van der Marel, J. H. van Boom, and M. Benziman. Regulation of cellulose synthesis in *Acetobacter xylinum* by cyclic diguanylic acid. *Nature* 325 (1987), pp. 279–281.

[4] U. Jenal and J. Malone. Mechanisms of cyclic-di-GMP signaling in bacteria. *Annu. Rev. Genet.* 40 (2006), pp. 385–407.

[5] A. J. Wolfe and K. L. Visick. Get the message out: cyclic-Di-GMP regulates multiple levels of flagellum-based motility. *J. Bacteriol.* 190 (2008), pp. 463–475.

[6] R. Hengge. Principles of c-di-GMP signalling in bacteria. *Nat. Rev. Microbiol.* 7 (2009), pp. 263–273.

[7] J. M. Dow, Y. Fouhy, J. F. Lucey, and R. P. Ryan. The HD-GYP domain, cyclic di-GMP signaling, and bacterial virulence to plants. *Mol. Plant Microbe Interact.* 19 (2006), pp. 1378–1384.

[8] P. A. Cotter and S. Stibitz. c-di-GMP-mediated regulation of virulence and biofilm formation. *Curr. Opin. Microbiol.* 10 (2007), pp. 17–23.

[9] R. P. Ryan, Y. Fouhy, J. F. Lucey, B. L. Jiang, Y. Q. He, J. X. Feng, J. L. Tang, and J. M. Dow. Cyclic di-GMP signalling in the virulence and environmental adaptation of *Xanthomonas campestris*. *Mol. Microbiol.* 63 (2007), pp. 429–442.

[10] A. Duerig, S. Abel, M. Folcher, M. Nicollier, T. Schwede, N. Amiot, B. Giese, and U. Jenal. Second messenger-mediated spatiotemporal control of protein degradation regulates bacterial cell cycle progression. *Genes Dev.* 23 (2009), pp. 93–104.

[11] P. C. Fineran, N. R. Williamson, K. S. Lilley, and G. P. Salmond. Virulence and prodigiosin antibiotic biosynthesis in *Serratia* are regulated pleiotropically by the GGDEF/EAL domain protein, PigX. *J. Bacteriol.* 189 (2007), pp. 7653–7662.

[12] M. Egli, R. V. Gessner, L. D. Williams, G. J. Quigley, G. A. van der Marel, J. H. van Boom, A. Rich, and C. A. Frederick. Atomic-resolution structure of the cellulose synthase regulator cyclic diguanylic acid. *Proc. Natl. Acad. Sci. U.S.A.* 87 (1990), pp. 3235–3239.

[13] Y. C. Liaw, Y. G. Gao, H. Robinson, G. M. Sheldrick, L. A. Sliedregt, G. A. van der Marel, J. H. van Boom, and A. H. Wang. Cyclic diguanylic acid behaves as a host molecule for planar intercalators. *FEBS Lett.* 264 (1990), pp. 223–227.

[14] Y. Guan, Y. G. Gao, Y. C. Liaw, H. Robinson, and A. H. Wang. Molecular structure of cyclic diguanylic acid at 1 Å resolution of two crystal forms: self-association, interactions with metal ion/planar dyes and modeling studies. *J. Biomol. Struct. Dyn.* 11 (1993), pp. 253–276.

[15] C. Chan, R. Paul, D. Samoray, N. C. Amiot, B. Giese, U. Jenal, and T. Schirmer. Structural basis of activity and allosteric control of diguanylate cyclase. *Proc. Natl. Acad. Sci. U.S.A.* 101 (2004), pp. 17084–17089.

[16] P. Wassmann, C. Chan, R. Paul, A. Beck, H. Heerklotz, U. Jenal, and T. Schirmer. Structure of BeF_3^- -modified response regulator PleD: implications for diguanylate cyclase activation, catalysis, and feedback inhibition. *Structure* 15 (2007), pp. 915–927.

[17] N. De, M. Pirruccello, P. V. Krasteva, N. Bae, R. V. Raghavan, and H. Sondermann. Phosphorylation-independent regulation of the diguanylate cyclase WspR. *PLoS Biol.* 6 (2008), e67.

[18] N. De, M. V. Navarro, R. V. Raghavan, and H. Sondermann. Determinants for the activation and autoinhibition of the diguanylate cyclase response regulator WspR. *J. Mol. Biol.* 393 (2009), pp. 619–633.

[19] J. Ko, K. S. Ryu, H. Kim, J. S. Shin, J. O. Lee, C. Cheong, and B. S. Choi. Structure of PP4397 reveals the molecular basis for different c-di-GMP binding modes by PilZ domain proteins. *J. Mol. Biol.* 398 (2010), pp. 97–110.

[20] J. Habazettl, M. G. Allan, U. Jenal, and S. Grzesiek. Solution structure of the PilZ domain protein PA4608 complex with cyclic di-GMP identifies charge clustering as molecular readout. *J. Biol. Chem.* 286 (2011), pp. 14304–14314.

[21] P. V. Krasteva, J. C. Fong, N. J. Shikuma, S. Beyhan, M. V. Navarro, F. H. Yildiz, and H. Sondermann. Vibrio cholerae VpsT regulates matrix production and motility by directly sensing cyclic di-GMP. *Science* 327 (2010), pp. 866–868.

[22] G. Minasov, S. Padavattan, L. Shuvalova, J. S. Brunzelle, D. J. Miller, A. Basle, C. Massa, F. R. Collart, T. Schirmer, and W. F. Anderson. Crystal structures of YkuI and its complex with second messenger cyclic Di-GMP suggest catalytic mechanism of phosphodiester bond cleavage by EAL domains. *J. Biol. Chem.* 284 (2009), pp. 13174–13184.

[23] T. R. Barends, E. Hartmann, J. J. Griese, T. Beitlich, N. V. Kirienko, D. A. Ryjenkov, J. Reinstein, R. L. Shoeman, M. Gomelsky, and I. Schlichting. Structure and mechanism of a bacterial light-regulated cyclic nucleotide phosphodiesterase. *Nature* 459 (2009), pp. 1015–1018.

[24] M. V. Navarro, N. De, N. Bae, Q. Wang, and H. Sondermann. Structural analysis of the GGDEF-EAL domain-containing c-di-GMP receptor FimX. *Structure* 17 (2009), pp. 1104–1116.

[25] J. Benach, S. S. Swaminathan, R. Tamayo, S. K. Handelman, E. Folta-Stogniew, J. E. Ramos, F. Forouhar, H. Neely, J. Seetharaman, A. Camilli, and J. F. Hunt. The structural basis of cyclic diguanylate signal transduction by PilZ domains. *EMBO J.* 26 (2007), pp. 5153–5166.

[26] Z. Zhang, B. L. Gaffney, and R. A. Jones. C-di-GMP displays a monovalent metal ion-dependent polymorphism. *J. Am. Chem. Soc.* 126 (2004), pp. 16700–16701.

[27] Z. Zhang, S. Kim, B. L. Gaffney, and R. A. Jones. Polymorphism of the signaling molecule c-di-GMP. *J. Am. Chem. Soc.* 128 (2006), pp. 7015–7024.

[28] M. Gentner, M. G. Allan, F. Zaehringer, T. Schirmer, and S. Grzesiek. Oligomer formation of the bacterial second messenger c-di-GMP: reaction rates and equilibrium constants indicate a monomeric state at physiological concentrations. *J. Am. Chem. Soc.* 134 (2) (2011), pp. 1019–1029.

[29] R. Paul, S. Weiser, N. C. Amiot, C. Chan, T. Schirmer, B. Giese, and U. Jenal. Cell cycle-dependent dynamic localization of a bacterial response regulator with a novel di-guanylate cyclase output domain. *Genes Dev.* 18 (2004), pp. 715–727.

[30] J. G. Malone, R. Williams, M. Christen, U. Jenal, A. J. Spiers, and P. B. Rainey. The structure-function relationship of WspR, a *Pseudomonas fluorescens* response regulator with a GGDEF output domain. *Microbiology (Reading, Engl.)* 153 (2007), pp. 980–994.

[31] B. Christen, M. Christen, R. Paul, F. Schmid, M. Folcher, P. Jenoe, M. Meuwly, and U. Jenal. Allosteric control of cyclic di-GMP signaling. *J. Biol. Chem.* 281 (2006), pp. 32015–32024.

[32] M. Christen, B. Christen, M. Folcher, A. Schauerte, and U. Jenal. Identification and characterization of a cyclic di-GMP-specific phosphodiesterase and its allosteric control by GTP. *J. Biol. Chem.* 280 (2005), pp. 30829–30837.

[33] A. J. Schmidt, D. A. Ryjenkov, and M. Gomelsky. The ubiquitous protein domain EAL is a cyclic diguanylate-specific phosphodiesterase: enzymatically active and inactive EAL domains. *J. Bacteriol.* 187 (2005), pp. 4774–4781.

[34] R. Tamayo, A. D. Tischler, and A. Camilli. The EAL domain protein VieA is a cyclic diguanylate phosphodiesterase. *J. Biol. Chem.* 280 (2005), pp. 33324–33330.

[35] R. P. Ryan, Y. Fouhy, J. F. Lucey, L. C. Crossman, S. Spiro, Y. W. He, L. H. Zhang, S. Heeb, M. Camara, P. Williams, and J. M. Dow. Cell-cell signaling in *Xanthomonas campestris* involves an HD-GYP domain protein that functions in cyclic di-GMP turnover. *Proc. Natl. Acad. Sci. U.S.A.* 103 (2006), pp. 6712–6717.

[36] F. Rao, Y. Yang, Y. Qi, and Z. X. Liang. Catalytic mechanism of cyclic di-GMP-specific phosphodiesterase: a study of the EAL domain-containing RocR from *Pseudomonas aeruginosa*. *J. Bacteriol.* 190 (2008), pp. 3622–3631.

[37] M. Y. Galperin, D. A. Natale, L. Aravind, and E. V. Koonin. A specialized version of the HD hydrolase domain implicated in signal transduction. *J. Mol. Microbiol. Biotechnol.* 1 (1999), pp. 303–305.

[38] K. Suzuki, P. Babitzke, S. R. Kushner, and T. Romeo. Identification of a novel regulatory protein (CsrD) that targets the global regulatory RNAs CsrB and CsrC for degradation by RNase E. *Genes Dev.* 20 (2006), pp. 2605–2617.

[39] N. Tschowri, S. Busse, and R. Hengge. The BLUF-EAL protein YcgF acts as a direct anti-repressor in a blue-light response of *Escherichia coli*. *Genes Dev.* 23 (2009), pp. 522–534.

[40] A. S. Seshasayee, G. M. Fraser, and N. M. Luscombe. Comparative genomics of cyclic-di-GMP signalling in bacteria: post-translational regulation and catalytic activity. *Nucleic Acids Res.* 38 (2010), pp. 5970–5981.

[41] L. E. Ulrich, E. V. Koonin, and I. B. Zhulin. One-component systems dominate signal transduction in prokaryotes. *Trends Microbiol.* 13.2 (2005), pp. 52–56.

[42] J. R. Tuckerman, G. Gonzalez, E. H. Sousa, X. Wan, J. A. Saito, M. Alam, and M. A. Gilles-Gonzalez. An oxygen-sensing diguanylate cyclase and phosphodiesterase couple for c-di-GMP control. *Biochemistry* 48 (2009), pp. 9764–9774.

[43] J. R. Tuckerman, G. Gonzalez, and M. A. Gilles-Gonzalez. Cyclic di-GMP activation of polynucleotide phosphorylase signal-dependent RNA processing. *J. Mol. Biol.* 407 (2011), pp. 633–639.

[44] Y. Qi, F. Rao, Z. Luo, and Z. X. Liang. A flavin cofactor-binding PAS domain regulates c-di-GMP synthesis in AxDGC2 from *Acetobacter xylinum*. *Biochemistry* 48 (2009), pp. 10275–10285.

[45] H. K. Carlson, R. E. Vance, and M. A. Marletta. H-NOX regulation of c-di-GMP metabolism and biofilm formation in *Legionella pneumophila*. *Mol. Microbiol.* 77 (2010), pp. 930–942.

[46] R. Paul, S. Abel, P. Wassmann, A. Beck, H. Heerklotz, and U. Jenal. Activation of the diguanylate cyclase PleD by phosphorylation-mediated dimerization. *J. Biol. Chem.* 282 (2007), pp. 29170–29177.

[47] A. Boehm, S. Steiner, F. Zaehringer, A. Casanova, F. Hamburger, D. Ritz, W. Keck, M. Ackermann, T. Schirmer, and U. Jenal. Second messenger signalling governs *Escherichia coli* biofilm induction upon ribosomal stress. *Mol. Microbiol.* 72 (2009), pp. 1500–1516.

[48] X. Wang, J. F. Preston, and T. Romeo. The *pgaABCD* locus of *Escherichia coli* promotes the synthesis of a polysaccharide adhesin required for biofilm formation. *J. Bacteriol.* 186 (2004), pp. 2724–2734.

[49] K. Jonas, A. N. Edwards, R. Simm, T. Romeo, U. Romling, and O. Melefors. The RNA binding protein CsrA controls cyclic di-GMP metabolism by directly regulating the expression of GGDEF proteins. *Mol. Microbiol.* 70 (2008), pp. 236–257.

[50] X. Wang, A. K. Dubey, K. Suzuki, C. S. Baker, P. Babitzke, and T. Romeo. CsrA post-transcriptionally represses *pgaABCD*, responsible for synthesis of a biofilm polysaccharide adhesin of *Escherichia coli*. *Mol. Microbiol.* 56 (2005), pp. 1648–1663.

[51] R. D. Finn, J. Mistry, J. Tate, P. Coggill, A. Heger, J. E. Pollington, O. L. Gavin, P. Gunasekaran, G. Ceric, K. Forslund, L. Holm, E. L. Sonnhammer, S. R. Eddy, and A. Bateman. The Pfam protein families database. *Nucleic Acids Res.* 38 (2010), pp. D211–222.

[52] J. Draper, K. Karplus, and K. M. Ottemann. Identification of a chemoreceptor zinc-binding domain common to cytoplasmic bacterial chemoreceptors. *J. Bacteriol.* 193 (2011), pp. 4338–4345.

[53] B. L. Vallee and K. H. Falchuk. The biochemical basis of zinc physiology. *Physiol. Rev.* 73 (1993), pp. 79–118.

[54] C. E. Outten, D. A. Tobin, J. E. Penner-Hahn, and T. V. O'Halloran. Characterization of the metal receptor sites in *Escherichia coli* Zur, an ultrasensitive zinc(II) metalloregulatory protein. *Biochemistry* 40 (2001), pp. 10417–10423.

[55] K. Hantke. Bacterial zinc transporters and regulators. *Biometals* 14 (2001), pp. 239–249.

[56] D. S. Auld. Zinc coordination sphere in biochemical zinc sites. *Biometals* 14 (2001), pp. 271–313.

[57] W. Maret. Zinc coordination environments in proteins determine zinc functions. *J Trace Elem Med Biol* 19 (2005), pp. 7–12.

[58] S. Leonhartsberger, A. Huber, F. Lottspeich, and A. Bock. The *hydH/G* Genes from *Escherichia coli* code for a zinc and lead responsive two-component regulatory system. *J. Mol. Biol.* 307 (2001), pp. 93–105.

[59] K. R. Brocklehurst, J. L. Hobman, B. Lawley, L. Blank, S. J. Marshall, N. L. Brown, and A. P. Morby. ZntR is a Zn(II)-responsive MerR-like transcriptional regulator of zntA in *Escherichia coli*. *Mol. Microbiol.* 31 (1999), pp. 893–902.

[60] A. I. Graham, S. Hunt, S. L. Stokes, N. Bramall, J. Bunch, A. G. Cox, C. W. McLeod, and R. K. Poole. Severe zinc depletion of *Escherichia coli*: roles for high affinity zinc binding by ZinT, zinc transport and zinc-independent proteins. *J. Biol. Chem.* 284 (2009), pp. 18377–18389.

[61] S. I. Patzer and K. Hantke. The ZnuABC high-affinity zinc uptake system and its regulator Zur in *Escherichia coli*. *Mol. Microbiol.* 28 (1998), pp. 1199–1210.

[62] A. Changela, K. Chen, Y. Xue, J. Holschen, C. E. Outten, T. V. O'Halloran, and A. Mondragon. Molecular basis of metal-ion selectivity and zeptomolar sensitivity by CueR. *Science* 301 (2003), pp. 1383–1387.

[63] D. Lucarelli, S. Russo, E. Garman, A. Milano, W. Meyer-Klaucke, and E. Pohl. Crystal structure and function of the zinc uptake regulator FurB from *Mycobacterium tuberculosis*. *J. Biol. Chem.* 282 (2007), pp. 9914–9922.

[64] Y. Hitomi, C. E. Outten, and T. V. O'Halloran. Extreme zinc-binding thermodynamics of the metal sensor/regulator protein ZntR. *J. Am. Chem. Soc.* 123 (2001), pp. 8614–8615.

[65] D. A. Ryjenkov, R. Simm, U. Romling, and M. Gomelsky. The PilZ domain is a receptor for the second messenger c-di-GMP: the PilZ domain protein YcgR controls motility in enterobacteria. *J. Biol. Chem.* 281 (2006), pp. 30310–30314.

[66] M. Christen, B. Christen, M. G. Allan, M. Folcher, P. Jeno, S. Grzesiek, and U. Jenal. DgrA is a member of a new family of cyclic diguanosine monophosphate receptors and controls flagellar motor function in *Caulobacter crescentus*. *Proc. Natl. Acad. Sci. U.S.A.* 104 (2007), pp. 4112–4117.

[67] J. T. Pratt, R. Tamayo, A. D. Tischler, and A. Camilli. PilZ domain proteins bind cyclic diguanylate and regulate diverse processes in *Vibrio cholerae*. *J. Biol. Chem.* 282 (2007), pp. 12860–12870.

[68] Y. McCarthy, R. P. Ryan, K. O'Donovan, Y. Q. He, B. L. Jiang, J. X. Feng, J. L. Tang, and J. M. Dow. The role of PilZ domain proteins in the virulence of *Xanthomonas campestris pv. campestris*. *Mol. Plant Pathol.* 9 (2008), pp. 819–824.

[69] J. W. Hickman and C. S. Harwood. Identification of FleQ from *Pseudomonas aeruginosa* as a c-di-GMP-responsive transcription factor. *Mol. Microbiol.* 69 (2008), pp. 376–389.

[70] V. T. Lee, J. M. Matewish, J. L. Kessler, M. Hyodo, Y. Hayakawa, and S. Lory. A cyclic-di-GMP receptor required for bacterial exopolysaccharide production. *Mol. Microbiol.* 65 (2007), pp. 1474–1484.

[71] S. Beyhan, L. S. Odell, and F. H. Yildiz. Identification and characterization of cyclic diguanylate signaling systems controlling rugosity in *Vibrio cholerae*. *J. Bacteriol.* 190 (2008), pp. 7392–7405.

[72] N. Sudarsan, E. R. Lee, Z. Weinberg, R. H. Moy, J. N. Kim, K. H. Link, and R. R. Breaker. Riboswitches in eubacteria sense the second messenger cyclic di-GMP. *Science* 321 (2008), pp. 411–413.

[73] I. Sutherland. Biofilm exopolysaccharides: a strong and sticky framework. *Microbiology (Reading, Engl.)* 147 (2001), pp. 3–9.

[74] S. S. Branda, S. Vik, L. Friedman, and R. Kolter. Biofilms: the matrix revisited. *Trends Microbiol.* 13 (2005), pp. 20–26.

[75] X. Zogaj, M. Nimtz, M. Rohde, W. Bokranz, and U. Romling. The multicellular morphotypes of *Salmonella typhimurium* and *Escherichia coli* produce cellulose as the second component of the extracellular matrix. *Mol. Microbiol.* 39 (2001), pp. 1452–1463.

[76] L. Friedman and R. Kolter. Two genetic loci produce distinct carbohydrate-rich structural components of the *Pseudomonas aeruginosa* biofilm matrix. *J. Bacteriol.* 186 (2004), pp. 4457–4465.

[77] L. Ma, H. Lu, A. Sprinkle, M. R. Parsek, and D. J. Wozniak. *Pseudomonas aeruginosa* Psl is a galactose- and mannose-rich exopolysaccharide. *J. Bacteriol.* 189 (2007), pp. 8353–8356.

[78] D. Mack, W. Fischer, A. Krokotsch, K. Leopold, R. Hartmann, H. Egge, and R. Laufs. The intercellular adhesin involved in biofilm accumulation of *Staphylococcus epidermidis* is a linear β-1,6-linked glucosaminoglycan: purification and structural analysis. *J. Bacteriol.* 178 (1996), pp. 175–183.

[79] K. Agladze, D. Jackson, and T. Romeo. Periodicity of cell attachment patterns during *Escherichia coli* biofilm development. *J. Bacteriol.* 185 (2003), pp. 5632–5638.

[80] K. Agladze, X. Wang, and T. Romeo. Spatial periodicity of *Escherichia coli* K-12 biofilm microstructure initiates during a reversible, polar attachment phase of development and requires the polysaccharide adhesin PGA. *J. Bacteriol.* 187 (2005), pp. 8237–8246.

[81] C. O. Jarrett, E. Deak, K. E. Isherwood, P. C. Oyston, E. R. Fischer, A. R. Whitney, S. D. Kobayashi, F. R. DeLeo, and B. J. Hinnebusch. Transmission of *Yersinia pestis* from an infectious biofilm in the flea vector. *J. Infect. Dis.* 190 (2004), pp. 783–792.

[82] C. Vuong, J. M. Voyich, E. R. Fischer, K. R. Braughton, A. R. Whitney, F. R. DeLeo, and M. Otto. Polysaccharide intercellular adhesin (PIA) protects *Staphylococcus epidermidis* against major components of the human innate immune system. *Cell. Microbiol.* 6 (2004), pp. 269–275.

[83] J. Begun, J. M. Gaiani, H. Rohde, D. Mack, S. B. Calderwood, F. M. Ausubel, and C. D. Sifri. Staphylococcal biofilm exopolysaccharide protects against *Caenorhabditis elegans* immune defenses. *PLoS Pathog.* 3 (2007), e57.

[84] E. A. Izano, I. Sadovskaya, E. Vinogradov, M. H. Mulks, K. Velliyagounder, C. Ragunath, W. B. Kher, N. Ramasubbu, S. Jabbouri, M. B. Perry, and J. B. Kaplan. Poly-N-acetylglucosamine mediates biofilm formation and antibiotic resistance in *Actinobacillus pleuropneumoniae*. *Microb. Pathog.* 43 (2007), pp. 1–9.

[85] D. W. Jackson, K. Suzuki, L. Oakford, J. W. Simecka, M. E. Hart, and T. Romeo. Biofilm formation and dispersal under the influence of the global regulator CsrA of *Escherichia coli*. *J. Bacteriol.* 184 (2002), pp. 290–301.

[86] C. Goller, X. Wang, Y. Itoh, and T. Romeo. The cation-responsive protein NhaR of *Escherichia coli* activates *pgaABCD* transcription, required for production of the biofilm adhesin poly-β-1,6-N-acetyl-D-glucosamine. *J. Bacteriol.* 188 (2006), pp. 8022–8032.

[87] U. M. Unligil and J. M. Rini. Glycosyltransferase structure and mechanism. *Curr. Opin. Struct. Biol.* 10 (2000), pp. 510–517.

[88] Y. Itoh, J. D. Rice, C. Goller, A. Pannuri, J. Taylor, J. Meisner, T. J. Beveridge, J. F. Preston, and T. Romeo. Roles of *pgaABCD* genes in synthesis, modification, and export of the *Escherichia coli* biofilm adhesin poly-β-1,6-N-acetyl-D-glucosamine. *J. Bacteriol.* 190 (2008), pp. 3670–3680.

[89] S. Forman, A. G. Bobrov, O. Kirillina, S. K. Craig, J. Abney, J. D. Fetherston, and R. D. Perry. Identification of critical amino acid residues in the plague biofilm Hms proteins. *Microbiology (Reading, Engl.)* 152 (2006), pp. 3399–3410.

[90] I. G. Boneca. The role of peptidoglycan in pathogenesis. *Curr. Opin. Microbiol.* 8 (2005), pp. 46–53.

[91] A. Severin, K. Tabei, and A. Tomasz. The structure of the cell wall peptidoglycan of *Bacillus cereus* RSVF1, a strain closely related to *Bacillus anthracis*. *Microb. Drug Resist.* 10 (2004), pp. 77–82.

[92] A. Bera, S. Herbert, A. Jakob, W. Vollmer, and F. Gotz. Why are pathogenic staphylococci so lysozyme resistant? The peptidoglycan O-acetyltransferase OatA is the major determinant for lysozyme resistance of *Staphylococcus aureus*. *Mol. Microbiol.* 55 (2005), pp. 778–787.

[93] D. E. Blair and D. M. van Aalten. Structures of *Bacillus subtilis* PdaA, a family 4 carbohydrate esterase, and a complex with N-acetyl-glucosamine. *FEBS Lett.* 570 (2004), pp. 13–19.

[94] F. Caufrier, A. Martinou, C. Dupont, and V. Bouriotis. Carbohydrate esterase family 4 enzymes: substrate specificity. *Carbohydr. Res.* 338 (2003), pp. 687–692.

[95] A. Martinou, D. Koutsioulis, and V. Bouriotis. Expression, purification, and characterization of a cobalt-activated chitin deacetylase (Cda2p) from *Saccharomyces cerevisiae*. *Protein Expr. Purif.* 24 (2002), pp. 111–116.

[96] R. Kawai, R. Nagata, A. Hirata, and Y. Hayakawa. A new synthetic approach to cyclic bis(3'→5')diguanylic acid. *Nucleic Acids Res. Suppl.* 3 (2003), pp. 103–104.

[97] Y. Hayakawa, R. Nagata, A. Hirata, M. Hyodo, and R. Kawai. A facile synthetis of cyclic bis(3'-5')diguanylic acid. *Tetrahedron* 59 (2003), pp. 6465–6471.

[98] N. Amiot, K. Heintz, and B. Giese. New Approach for the Synthesis of c-di-GMP and Its Analogues. *Synthesis* 24 (2006), pp. 4230–4236.

[99] H. Yan and A. L. Aguilar. Synthesis of 3',5'-cyclic diguanylic acid (cdiGMP) using 1-(4-chlorophenyl)-4-ethoxypiperidin-4-yl as a protecting group for 2'-hydroxy functions of ribonucleosides. *Nucleosides Nucleotides Nucleic Acids* 26 (2007), pp. 189–204.

[100] I. Kiburu, A. Shurer, L. Yan, and H. O. Sintim. A simple solid-phase synthesis of the ubiquitous bacterial signaling molecule, c-di-GMP and analogues. *Mol Biosyst* 4 (2008), pp. 518–520.

[101] B. I. Kazmierczak, M. B. Lebron, and T. S. Murray. Analysis of FimX, a phosphodiesterase that governs twitching motility in *Pseudomonas aeruginosa*. *Mol. Microbiol.* 60 (2006), pp. 1026–1043.

[102] M. Merighi, V. T. Lee, M. Hyodo, Y. Hayakawa, and S. Lory. The second messenger bis-(3'-5')-cyclic-GMP and its PilZ domain-containing receptor Alg44 are required for alginate biosynthesis in *Pseudomonas aeruginosa*. *Mol. Microbiol.* 65 (2007), pp. 876–895.

[103] F. Rao, S. Pasunooti, Y. Ng, W. Zhuo, L. Lim, A. W. Liu, and Z. X. Liang. Enzymatic synthesis of c-di-GMP using a thermophilic diguanylate cyclase. *Anal. Biochem.* 389 (2009), pp. 138–142.

[104] W. A. Hendrickson, J. R. Horton, and D. M. LeMaster. Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J.* 9 (1990), pp. 1665–1672.

[105] J. B. Hunt, S. H. Neece, and A. Ginsburg. The use of 4-(2-pyridylazo)resorcinol in studies of zinc release from *Escherichia coli* aspartate transcarbamoylase. *Anal. Biochem.* 146 (1985), pp. 150–157.

[106] A. A. Baykov, O. A. Evtushenko, and S. M. Avaeva. A malachite green procedure for orthophosphate determination and its use in alkaline phosphatase-based enzyme immunoassay. *Anal. Biochem.* 171 (1988), pp. 266–270.

[107] Z. X. Wang. An exact mathematical expression for describing competitive binding of two different ligands to a protein molecule. *FEBS Lett.* 360 (1995), pp. 111–114.

[108] W. Kabsch. XDS. *Acta Crystallogr. D Biol. Crystallogr.* 66 (2010), pp. 125–132.

[109] A. G. Leslie. The integration of macromolecular diffraction data. *Acta Crystallogr. D Biol. Crystallogr.* 62 (2006), pp. 48–57.

[110] P. Evans. Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* 62 (2006), pp. 72–82.

[111] C. Vonrhein, E. Blanc, P. Roversi, and G. Bricogne. Automated structure solution with autoSHARP. *Methods Mol. Biol.* 364 (2007), pp. 215–230.

[112] G. Langer, S. X. Cohen, V. S. Lamzin, and A. Perrakis. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* 3 (2008), pp. 1171–1179.

[113] P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66 (2010), pp. 486–501.

[114] G. N. Murshudov, A. A. Vagin, and E. J. Dodson. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* 53 (1997), pp. 240–255.

[115] V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, and D. C. Richardson. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 66 (2010), pp. 12–21.

[116] S. C. Lovell, I. W. Davis, W. B. Arendall, P. I. de Bakker, J. M. Word, M. G. Prisant, J. S. Richardson, and D. C. Richardson. Structure validation by Cα geometry: ϕ, ψ and Cβ deviation. *Proteins* 50 (2003), pp. 437–450.

[117] T. C. Terwilliger, R. W. Grosse-Kunstleve, P. V. Afonine, N. W. Moriarty, P. H. Zwart, L. W. Hung, R. J. Read, and P. D. Adams. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr. D Biol. Crystallogr.* 64 (2008), pp. 61–69.

[118] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, and R. J. Read. Phaser crystallographic software. *J Appl Crystallogr* 40 (2007), pp. 658–674.

[119] E. Blanc, P. Roversi, C. Vonrhein, C. Flensburg, S. M. Lea, and G. Bricogne. Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallogr. D Biol. Crystallogr.* 60 (2004), pp. 2210–2221.

[120] W. Kabsch and C. Sander. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22 (1983), pp. 2577–2637.

[121] M. Heinig and D. Frishman. STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res.* 32 (2004), W500–502.

[122] E. Krissinel and K. Henrick. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372 (2007), pp. 774–797.

[123] E. Krissinel and K. Henrick. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D Biol. Crystallogr.* 60 (2004), pp. 2256–2268.

[124] Number 4 Collaborative Computational Project. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 50 (1994), pp. 760–763.

[125] L. Holm and P. Rosenstrom. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* 38 (2010), W545–549.

[126] A. Golovin and K. Henrick. MSDmotif: exploring protein sites and motifs. *BMC Bioinformatics* 9 (2008), p. 312.

[127] A. Hijikata, K. Yura, T. Noguti, and M. Go. Revisiting gap locations in amino acid sequence alignments and a proposal for a method to improve them by introducing solvent accessibility. *Proteins* 79 (2011), pp. 1868–1877.

[128] G. E. Crooks, G. Hon, J. M. Chandonia, and S. E. Brenner. WebLogo: a sequence logo generator. *Genome Res.* 14 (2004), pp. 1188–1190.

[129] C. Deprez, R. Lloubes, M. Gavioli, D. Marion, F. Guerlesquin, and L. Blanchard. Solution structure of the *E.coli* TolA C-terminal domain reveals conformational changes upon binding to the phage g3p N-terminal domain. *J. Mol. Biol.* 346 (2005), pp. 1047–1057.

[130] N. Eswar, B. Webb, M. A. Marti-Renom, M. S. Madhusudhan, D. Eramian, M. Y. Shen, U. Pieper, and A. Sali. Comparative protein structure modeling using Modeller. *Curr Protoc Bioinformatics* Chapter 5 (2006), Unit 5.6.

[131] Z. S. Derewenda. Rational protein crystallization by mutational surface engineering. *Structure* 12 (2004), pp. 529–535.

[132] Z. S. Derewenda and P. G. Vekilov. Entropy and surface engineering in protein crystallization. *Acta Crystallogr. D Biol. Crystallogr.* 62 (2006), pp. 116–124.

[133] L. Goldschmidt, D. R. Cooper, Z. S. Derewenda, and D. Eisenberg. Toward rational protein crystallization: A Web server for the design of crystallizable protein variants. *Protein Sci.* 16 (2007), pp. 1569–1576.

[134] A. P. Golovanov, G. M. Hautbergue, S. A. Wilson, and L. Y. Lian. A simple method for improving protein solubility and long-term stability. *J. Am. Chem. Soc.* 126 (2004), pp. 8933–8939.

[135] T. S. Walter, C. Meier, R. Assenberg, K. F. Au, J. Ren, A. Verma, J. E. Nettleship, R. J. Owens, D. I. Stuart, and J. M. Grimes. Lysine methylation as a routine rescue strategy for protein crystallization. *Structure* 14 (2006), pp. 1617–1622.

[136] A. Vagin and A. Teplyakov. Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* 66 (2010), pp. 22–25.

[137] I. L. Alberts, K. Nadassy, and S. J. Wodak. Analysis of zinc binding sites in protein crystal structures. *Protein Sci.* 7 (1998), pp. 1700–1716.

[138] J. W. Becker, A. I. Marcy, L. L. Rokosz, M. G. Axel, J. J. Burbaum, P. M. Fitzgerald, P. M. Cameron, C. K. Esser, W. K. Hagmann, and J. D. Hermes. Stromelysin-1: three-dimensional structure of the inhibited catalytic domain and of the C-truncated proenzyme. *Protein Sci.* 4 (1995), pp. 1966–1976.

[139] E. Morgunova, A. Tuuttila, U. Bergmann, M. Isupov, Y. Lindqvist, G. Schneider, and K. Tryggvason. Structure of human pro-matrix metalloproteinase-2: activation mechanism revealed. *Science* 284 (1999), pp. 1667–1670.

[140] P. A. Elkins, Y. S. Ho, W. W. Smith, C. A. Janson, K. J. D'Alessio, M. S. McQueney, M. D. Cummings, and A. M. Romanic. Structure of the C-terminally truncated human ProMMP9, a gelatin-binding matrix metalloproteinase. *Acta Crystallogr. D Biol. Crystallogr.* 58 (2002), pp. 1182–1192.

[141] D. Jozic, G. Bourenkov, N. H. Lim, R. Visse, H. Nagase, W. Bode, and K. Maskos. X-ray structure of human proMMP-1: new insights into procollagenase activation and collagen binding. *J. Biol. Chem.* 280 (2005), pp. 9578–9585.

[142] A. Mac Sweeney, S. Gil-Parrado, D. Vinzenz, A. Bernardi, A. Hein, U. Bodendorf, P. Erbel, C. Logel, and B. Gerhartz. Structural basis for the substrate specificity of bone morphogenetic protein 1/tolloid-like metalloproteases. *J. Mol. Biol.* 384 (2008), pp. 228–239.

[143] A. Torres-Larios, R. Sankaranarayanan, B. Rees, A. C. Dock-Bregeon, and D. Moras. Conformational movements and cooperativity upon amino acid, ATP and tRNA binding in threonyl-tRNA synthetase. *J. Mol. Biol.* 331 (2003), pp. 201–211.

[144] J. Ishijima, Y. Uchida, C. Kuroishi, C. Tuzuki, N. Takahashi, N. Okazaki, K. Yutani, and M. Miyano. Crystal structure of alanyl-tRNA synthetase editing-domain homolog (PH0574) from a hyperthermophile, *Pyrococcus horikoshii* OT3 at 1.45 A resolution. *Proteins* 62 (2006), pp. 1133–1137.

[145] M. Naganuma, S. Sekine, R. Fukunaga, and S. Yokoyama. Unique protein architecture of alanyl-tRNA synthetase for aminoacylation, editing, and dimerization. *Proc. Natl. Acad. Sci. U.S.A.* 106 (2009), pp. 8489–8494.

[146] Y. Sato, A. Yoshikawa, A. Yamagata, H. Mimura, M. Yamashita, K. Ookata, O. Nureki, K. Iwai, M. Komada, and S. Fukai. Structural basis for specific cleavage of Lys 63-linked polyubiquitin chains. *Nature* 455 (2008), pp. 358–362.

[147] K. Briknarova, C. J. Thomas, J. York, and J. H. Nunberg. Structure of a zinc-binding domain in the Junin virus envelope glycoprotein. *J. Biol. Chem.* 286 (2011), pp. 1528–1536.

[148] J. A. Ippolito and D. W. Christianson. Structure of an engineered His3Cys zinc binding site in human carbonic anhydrase II. *Biochemistry* 32 (1993), pp. 9901–9905.

[149] L. E. Ulrich and I. B. Zhulin. Four-helix bundle: a ubiquitous sensory module in prokaryotic signal transduction. *Bioinformatics* 21 Suppl 3 (2005), pp. i45–48.

[150] C. S. Bond. TopDraw: a sketchpad for protein structure topology cartoons. *Bioinformatics* 19 (2003), pp. 311–312.

[151] C. Y. Yang, K. H. Chin, M. L. Chuah, Z. X. Liang, A. H. Wang, and S. H. Chou. The structure and inhibition of a GGDEF diguanylate cyclase complexed with (c-di-GMP)(2) at the active site. *Acta Crystallogr. D Biol. Crystallogr.* 67.Pt 12 (2011), pp. 997–1008.

[152] G. Anderegg. *Critical Survey of Stability constants of EDTA Complexes (IUPAC Chemical Data Series No.14)*. Pergamon Press, 1977.

[153] U. Heinz, M. Kiefer, A. Tholey, and H. W. Adolph. On the competition for available zinc. *J. Biol. Chem.* 280 (2005), pp. 3197–3207.

[154] X. Wan, J. R. Tuckerman, J. A. Saito, T. A. Freitas, J. S. Newhouse, J. R. Denery, M. Y. Galperin, G. Gonzalez, M. A. Gilles-Gonzalez, and M. Alam. Globins synthesize the second messenger bis-(3'-5')-cyclic diguanosine monophosphate in bacteria. *J. Mol. Biol.* 388 (2009), pp. 262–270.

[155] D. E. Blair, A. W. Schuttelkopf, J. I. MacRae, and D. M. van Aalten. Structure and metal-dependent mechanism of peptidoglycan deacetylase, a streptococcal virulence factor. *Proc. Natl. Acad. Sci. U.S.A.* 102 (2005), pp. 15429–15434.

[156] H.-O. Beutler. *Determination with Acetyl-CoA Synthetase. In Methods of Enzymatic Analysis.* Ed. by Bergmeyer H. U. VCH Publishers(UK) Ltd., Cambridge, UK, 1988.

[157] J. M. Frere. β-Lactamases and bacterial resistance to antibiotics. *Mol. Microbiol.* 16 (1995), pp. 385–395.

[158] A. A. Medeiros. Evolution and dissemination of β-lactamases accelerated by generations of β-lactam antibiotics. *Clin. Infect. Dis.* 24 Suppl 1 (1997), pp. 19–45.

[159] R. P. Ambler. The structure of β-lactamases. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 289 (1980), pp. 321–331.

[160] P. Huovinen, S. Huovinen, and G. A. Jacoby. Sequence of PSE-2 β-lactamase. *Antimicrob. Agents Chemother.* 32 (1988), pp. 134–136.

[161] B. Jaurin and T. Grundstrom. *ampC* cephalosporinase of *Escherichia coli* K-12 has a different evolutionary origin from that of β-lactamases of the penicillinase type. *Proc. Natl. Acad. Sci. U.S.A.* 78 (1981), pp. 4897–4901.

[162] A. Felici, G. Amicosante, A. Oratore, R. Strom, P. Ledent, B. Joris, L. Fanuel, and J. M. Frere. An overview of the kinetic parameters of class B β-lactamases. *Biochem. J.* 291 (Pt 1) (1993), pp. 151–155.

[163] J. M. Frere, B. Joris, B. Granier, A. Matagne, F. Jacob, and C. Bourguignon-Bellefroid. Diversity of the mechanisms of resistance to β-lactam antibiotics. *Res. Microbiol.* 142 (1991), pp. 705–710.

[164] J. Fisher, J. G. Belasco, S. Khosla, and J. R. Knowles. β-Lactamase proceeds via an acyl-enzyme intermediate. Interaction of the *Escherichia coli* RTEM enzyme with cefoxitin. *Biochemistry* 19 (1980), pp. 2895–2901.

[165] P. Swaren, L. Maveyraud, X. Raquet, S. Cabantous, C. Duez, J. D. Pedelacq, S. Mariotte-Boyer, L. Mourey, R. Labia, M. H. Nicolas-Chanoine, P. Nordmann, J. M. Frere, and J. P. Samama. X-ray analysis of the NMC-A β-lactamase at 1.64 Å resolution, a class A carbapenemase with broad substrate specificity. *J. Biol. Chem.* 273 (1998), pp. 26714–26721.

[166] A. P. Kuzin, M. Nukaga, Y. Nukaga, A. M. Hujer, R. A. Bonomo, and J. R. Knox. Structure of the SHV-1 β-lactamase. *Biochemistry* 38 (1999), pp. 5720–5727.

[167] W. Sougakoff, G. L'Hermite, L. Pernot, T. Naas, V. Guillet, P. Nordmann, V. Jarlier, and J. Delettre. Structure of the imipenem-hydrolyzing class A β-lactamase SME-1 from *Serratia marcescens*. *Acta Crystallogr. D Biol. Crystallogr.* 58 (2002), pp. 267–274.

[168] C. A. Smith, M. Caccamo, K. A. Kantardjieff, and S. Vakulenko. Structure of GES-1 at atomic resolution: insights into the evolution of carbapenamase activity in the class A extended-spectrum β-lactamases. *Acta Crystallogr. D Biol. Crystallogr.* 63 (2007), pp. 982–992.

[169] W. Ke, C. R. Bethel, J. M. Thomson, R. A. Bonomo, and F. van den Akker. Crystal structure of KPC-2: insights into carbapenemase activity in class A β-lactamases. *Biochemistry* 46 (2007), pp. 5732–5740.

[170] J. Walther-Rasmussen and N. H∅iby. Class A carbapenemases. *J. Antimicrob. Chemother.* 60 (2007), pp. 470–482.

[171] A. M. Queenan and K. Bush. Carbapenemases: the versatile β-lactamases. *Clin. Microbiol. Rev.* 20 (2007), pp. 440–458.

[172] D. M. Livermore and N. Woodford. The β-lactamase threat in *Enterobacteriaceae*, *Pseudomonas* and *Acinetobacter*. *Trends Microbiol.* 14 (2006), pp. 413–420.

[173] L. K. Siu. Antibiotics: action and resistance in gram-negative bacteria. *J Microbiol Immunol Infect* 35 (2002), pp. 1–11.

[174] K. Poole. Resistance to β-lactam antibiotics. *Cell. Mol. Life Sci.* 61 (2004), pp. 2200–2223.

[175] M. S. Helfand and R. A. Bonomo. β-lactamases: a survey of protein diversity. *Curr Drug Targets Infect Disord* 3 (2003), pp. 9–23.

[176] K. Bush. The impact of β-lactamases on the development of novel antimicrobial agents. *Curr Opin Investig Drugs* 3 (2002), pp. 1284–1290.

[177] J. D. Buynak. Understanding the longevity of the β-lactam antibiotics and of antibiotic/β-lactamase inhibitor combinations. *Biochem. Pharmacol.* 71 (2006), pp. 930–940.

[178] K. Coleman. Diazabicyclooctanes (DBOs): a potent new class of non-*beta*-lactam*beta*-lactamase inhibitors. *Curr. Opin. Microbiol.* 14.5 (2011), pp. 550–555.

[179] T. O. Yeates. Detecting and overcoming crystal twinning. *Meth. Enzymol.* 276 (1997), pp. 344–358.

[180] A. Vagin and A. Teplyakov. An approach to multi-copy search in molecular replacement. *Acta Crystallogr. D Biol. Crystallogr.* 56 (2000), pp. 1622–1624.

[181] A. W. Schuttelkopf and D. M. van Aalten. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. D Biol. Crystallogr.* 60 (2004), pp. 1355–1363.

[182] A. Perrakis, R. Morris, and V. S. Lamzin. Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.* 6 (1999), pp. 458–463.

[183] R. A. Laskowski, M. W. MacArthur, D. S. Moss, and J. M. Thornton. *PROCHECK*: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* 26.2 (1993), pp. 283–291.

Acknowledgment

Zuerst möchte ich mich bei Prof. Tilman Schirmer bedanken für die Möglichkeit diese interessante Arbeit in seinem Labor durchzuführen, für seine Disskussionsbereitschaft und das Teilhaben an seinem kristallographischen Wissen.

Ich möchte mich bei Prof. Urs Jenal bedanken in dessen Kooperation diese Arbeit entstanden ist. Der immer wieder neue Ideen entwickelte, die das Projekt voran trieben.

Der Firma Basilea, insbesondere Stefan Reinelt, danke ich für die Möglichkeit Einblicke in das Structure based drug design zu gewinnen.

Ganz besonders Danken möchte ich Alex Böhm, der das YdeH-Projekt ins Leben rief und mit seiner unendlichen Begeisterung immer neue Motivation lieferte. Vielen Dank für das biologischem Hintergrundwissen, für die vielen Ideen und den Support auch aus der Ferne.

Ich möchte mich bei Samuel Seiner bedanken für die Zusammenarbeit im PgaB-Projekt, für das Bereitstellen des Substrates für den PgaB assay und für viele hilfreiche Diskussionen.

My special thanks goes to the past and present members of the Schirmer lab Aline Borer, Amit Sundriyal, Arnaud Goepfert, Camille Peitsch, Caroline Peneff, Cedric Hutter, Christophe Wirth, Claudia Massa, Dietrich Samoray, Frederic Stanger, Ludwig Zumthor, Marlise Frick, Nisha Vinayak, Paul Wassmann, Roman Lehener, Stefanie Kauer and Tillmann Heinisch for their helpfulness in the lab and for their funny discussions during the lunch breaks. My special thanks goes to Claudia Massa for all her support and to Christophe Wirth for the help in the starting time in Basel.

Ganz herzlich möchte ich mich bei Dietrich Samoray bedanken für die vielen Diskussionen und die Hilfe bei der Durchführung zahlreicher Experimente.

For prove reading I would like to thank Claudia Massa, Friederike Schmidt, Christophe Wirth and Tibor Dufner.

Bei Elvira Friedrich bedanke ich mich für die gemeinsamen Kaffee-Pausen und das Rheinschwimmen.

Ganz besonders bedanken möchte ich mich bei Georg Zocher, der für jedes kristallographische Problem eine Lösung hatte.

Danke an Riki, Feli und Theresa, für die vielen Ratschlägen das wissentschaftliche Leben betreffend, aber vor allem für die Freundschaft.

Meinen Eltern danke ich von ganzen Herzen für ihre Unterstützung und Motivation.

Ganz besonders möchte ich mich bei Tibor für seine Geduld und Unterstützung bedanken. Besonders dafür, dass du dafür gesorgt hast, dass auch nicht-wissenschaftliche Dinge ihre Wertigkeit nicht verloren haben.